中央大学博士論文

Bayesian decision making and its relation to frequentist methods

Masaaki Doi
土 居 正 明

博士（理学）

中央大学大学院
理 工 学 研 究 科
数 学 専 攻

平成29年度
２０１８年３月

Acknowledgements

First and foremost, I would like to express the deepest appreciation to my supervisor, Professor Sadanori Konishi who provided me insightful advice and continuous encouragement through the whole of my research and Ph.D student life. His support made it possible to finish this thesis.

I gratefully acknowledge the members of my Ph.D. dissertation committee: Professors Tokio Matsuyama, Fumitake Sakaori, Yasuo Ohashi, Toshinari Kamakura and Kunio Shimizu, for their time and valuable feedback on a preliminary version of this thesis.

I would also like to express my sincere gratitude to the co-authors of my papers: Professors Yohei Kawasaki, Kazuki Ide and Mr. Fumihiro Takahashi. Discussions with them are illuminating and the results are included in this thesis.

I would like to thank to Mr. Akio Mori, Kiyonobu Okada and Jun Atsumi, and other members in Clinical Data Science and Quality Management Department, Toray Industries, Inc. for many support and encouragement.

I would also like to thank to all Konishi Laboratory members, especially Professors Toshihiro Misumi, Heewon Park, Ibuki Hoshina and Mr. Kazuki Matsuda. Their sharp advice and discussions with them are reflected in this thesis.

I am very grateful to Professors Toshiro Tango, Kunihiko Takahashi, Masako Nishikawa, Kazue Yamaoka, Dr. Tetsuji Yokoyama and many great professors (including Professor Konishi) who gave me excellent lectures at National Institute of Public Health (NIPH). They taught me statistics and the fun of statistics. They and NIPH biostatistics & CMS RCT course Alumni with secretary Ms. Yoko Nezu are still encouraging and supporting me.

I would like to offer my special thanks to three more mathematicians: Professors Mitsuru Ikawa, Toshio Oshima, Toshiyuki Kobayashi. When I was an undergraduate and a master course student, they gave me a fundamental training in mathematics and showed me how great mathematicians think. Without them, I could not start and continue my research.

Finally, I would like to thank my mother Hideko Doi and late father Toshimitsu Doi for their love and encouragement. I cannot express in words the gratitude for them.

Contents

Chapter 1

Introduction

Statistical decision making is quite important for many fields, such as drug development, epidemiology, and other natural and social sciences. For almost all situations, frequentist methods have been mainly used. Even when we have some historical data, we can use them only at the planning stage. Recently, on the other hand, large databases have become available in many fields and there is an expectation that decision making utilizing the accumulated data will become more efficient. In line with this, Bayesian methods have been gaining attention.

In this thesis, we mainly focus on Bayesian evidence for decision making, especially "the posterior probabilities of some hypotheses being true". This type of Bayesian evidence has long been studied. For the binomial probability, Altham (1969) derived the exact expression for the posterior probability of the "odds ratio being less than 1" by using the cumulative distribution of the hypergeometric distribution and showed the relationship between this probability and the p-value of Fisher's exact test. Casella and Berger (1987) reconciled Bayesian and frequentist evidence for some general classes of distributions for the one group case. Recently, a guidance from US Food and Drug Administration (FDA) "Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials" (US Food and Drug Administration and others (2010)) stated that as the Bayesian hypothesis testing, we may use the posterior probability that a particular hypothesis is true, given the observed data. Now, this type of decision making is studied actively especially in biomedical fields. For example, Zaslavsky (2010) studied the Bayesian posterior probabilities of some hypotheses being true for the binomial probability and the Poisson rate parameter and compared them with frequentist p-values for the one group case. For a two group superiority test, Kawasaki and Miyaoka (2012b) expressed the posterior probability of superiority hypothesis being true for binomial probabilities by using the generalized hypergeometric series, and derived a normal approximation formula. Kawasaki and Miyaoka (2014) and Zaslavsky (2013) investigated the relationship between this probability and the p-value of Fisher's exact test. Kawasaki and Miyaoka (2012a) expressed the posterior probability of superiority hypothesis being true for Poisson rate parameters by using Gauss hypergeometric series, and derived a normal approximation formula.

For a non-inferiority test, Gamalo et al. (2016) and Ghosh et al. (2016) considered the normal mean for two groups and three groups (active, placebo, and active control), respectively. Kawasaki and Miyaoka (2013) expressed the posterior probability of non-inferiority hypothesis being true for binomial probabilities by generalized hypergeometric series. Zaslavsky (2013) derived an approximation formula for the posterior probability of non-inferiority hypothesis being true based on the risk difference and the exact formula for the posterior probability of non-inferiority hypothesis being true based on the risk ratio. Gamalo et al. (2011) investigated how to decide the non-inferiority margin based on historical trial data and evaluated the posterior probability by using normal approximations and Monte Carlo approximation. Kawasaki et al. (2016) considered the posterior probability of non-inferiority hypothesis being true for the Poisson rate parameters.

In this thesis, we mainly consider one-sided hypothesis testing. For the importance of one-sided test, Casella and Berger (1987) stated, "There is a direction of interest in many experiments, and saddling an experimenter with a two-sided test would not be appropriate." In addition, Zaslavsky (2013) stated, "From a practical perspective, one-sided hypotheses that aim for better performance or non-inferiority are very natural in the clinical environment".

More precisely, the theme for this thesis is in the following:

- Bayesian superiority and non-inferiority testing for Poisson rate parameters
- Bayesian superiority and equivalence testing for the variances of the normal distributions
- Bayesian non-inferiority test for binomial probabilities.

We study (i) the exact expressions for the posterior probabilities of some hypotheses being true, (ii) the situation where Bayesian and frequentist decisions coincide, and (iii) how much the prior distributions affect Bayesian evidence.

The remainder of this thesis is organized as follows:

- **In Chapter 2**, we briefly summarize the Bayesian method of decision making. Especially, we focus on the posterior probabilities of some hypotheses being true.

- **In Chapter 3**, we consider the Bayesian superiority hypothesis testing for Poisson rate parameters λ_1 and λ_2. We derive exact and quite simple expressions for the posterior probability of superiority hypothesis being true given the data X_1, X_2, i.e., $\Pr(\lambda_1 < \lambda_2 \mid X_1, X_2)$ and show the relationship between this posterior probability and the p-value of the frequentist conditional test. Then, we generalize the results to the posterior probability $\Pr(\lambda_1/\lambda_2 < c \mid X_1, X_2)$ and the corresponding p-value of the frequentist test.

- **In Chapter 4**, we consider the Bayesian non-inferiority hypothesis testing for Poisson rate parameters. We derive an exact and simple expression for $\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2)$ with non-inferiority margin Δ under mild conditions, and then construct a non-inferiority framework which can be considered as an extension of the conditional test stated in Chapter 3. In this framework, we can naturally treat switching from non-inferiority to superiority test.

- **In Chapter 5**, we consider the Bayesian superiority and equivalence testing for the variances σ_1^2 and σ_2^2 of two normal distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$. First, we treat the posterior probabiltity of superiority hypothesis being true given the data vecrots $\mathbf{x}_1$ and $\mathbf{x}_2$, i.e., $\Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$. We derive exact and quite simple expressions for this posterior probability. Then we show the relationship between $\Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$ and the p-value of the frequentist F-test where (i) the means μ_1, μ_2 are known and the priors of σ_1^2 and σ_2^2 are scaled inverse chi-square distributions, (ii) the means are unknown and the priors of μ_1, μ_2 are noninformative, and the priors of σ_1^2 and σ_2^2 are scaled inverse chi-square distributions, (iii) the means are unknown and the priors of (μ_1, σ_1^2) and (μ_2, σ_2^2) are normal-inverse-gamma distributions. We also evaluate the posterior probability of equivalence $\Pr(1/\Delta < \sigma_1/\sigma_2 < \Delta \mid \mathbf{x}_1, \mathbf{x}_2)$ which we are able to express quite simply.

- **In Chapter 6**, we consider the Bayesian non-inferiority hypothesis testing for binomial probabilities. We derive the exact expression for the posterior probability $\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)$ under mild conditions, and then consider a framework which can be interpreted as the Bayesian non-inferiority extension of Fisher's exact test. Finally, we calculate sample size when historical data can be utilized.

- **Chapter 7**, we presents some concluding remarks.

Chapter 2

Bayesian evidences for decision makings

In this chapter, we briefly overview several Bayesian evidences for decision makings, mainly based on the posterior probabilities of some hypotheses being true. First, we summarize the results of Casella and Berger (1987) which are stated in general form. Then, we summarize the results based on the specific distributions. Finally, we briefly comment on the Bayes factor.

2.1 Reconciling Bayesian and frequentist evidence in the one-sided hypothesis testing by Casella and Berger (1987)

We first consider the reconciliation between Bayesian and frequentist evidence shown by Casella and Berger (1987). Consider the testing hypotheses $H_0 : \theta \leq 0$ vs $H_1 : \theta > 0$ based on the observed data $X = x$. Here, we suppose that x has the probability density function $f(x - \theta)$ where $f(\cdot)$ is

(a) symmetric about zero.

(b) $f(x - \theta)$ has monotone likelihood ratio (MLR), that is, for any $\theta_1 < \theta_2$, the distribution $f(x - \theta_1)$ and $f(x - \theta_2)$ are distinct, and the ratio $f(x - \theta_2)/f(x - \theta_1)$ is a nondecreasing function of x.

For the frequentist perspective, p-value given $X = x$ is defined as

$$p(x) = \Pr(X \geq x \mid \theta = 0) = \int_x^\infty f(t)dt.$$

Next, consider the Bayesian posterior probability given $X = x$. Let the probability density function of the prior distribution of θ be $\pi(\theta)$. Then, the probability density function of the posterior distribution is

$$\pi(\theta \mid x) = \frac{f(x-\theta)\pi(\theta)}{\int_{-\infty}^{\infty} f(x-\theta)\pi(\theta)d\theta}.$$

Therefore, the probability of the hypothesis $H_0 : \theta \leq 0$ being true is

$$\begin{aligned}
\Pr(H_0 \mid x) &= \Pr(\theta \leq 0 \mid x) \\
&= \int_{-\infty}^{0} \pi(\theta \mid x)d\theta \\
&= \frac{\int_{-\infty}^{0} f(x-\theta)\pi(\theta)d\theta}{\int_{-\infty}^{\infty} f(x-\theta)\pi(\theta)d\theta}.
\end{aligned}$$

Here, Casella and Berger (1987) derived the following relationships between the Bayesian and frequentist evidences.

Theorem 2.1 (Casella and Berger (1987))**.** For the hypotheses $H_0 : \theta \leq 0$ vs $H_1 : \theta > 0$, if f is symmetric and has MLR and if $x > 0$, then

$$\inf_{\pi \in \Gamma_S} \Pr(H_0 \mid x) = \inf_{\pi \in \Gamma_{2PS}} \Pr(H_0 \mid x) \leq p(x)$$

where $\Gamma_S = \{$ all distributions symmetric about zero $\}$ and $\Gamma_{2PS} = \{$ all two-point distributions symmetric about 0 $\}$.

Theorem 2.2 (Casella and Berger (1987))**.** For the hypotheses $H_0 : \theta \leq 0$ vs $H_1 : \theta > 0$, if f is symmetric and has MLR and if $x > 0$, then

$$\inf_{\pi \in \Gamma_{US}} \Pr(H_0 \mid x) = \inf_{\pi \in U_S} \Pr(H_0 \mid x) = p(x)$$

where $\Gamma_{US} = \{$ all distributions with unimodal densities, symmetric about zero $\}$ and $U_S = \{$ all symmetric uniform distributions $\}$.

These theorems clarified some relationships between Bayesian posterior probability and frequentist p-values. However, the limitations are the following:

(i) Only one distribution (one group situation) is considered.

(ii) Especially for Theorem 2.1, the condition where the equality holds is not sufficiently characterized.

(iii) The assumption where f is symmetric about 0 is quite strong.

In this thesis,

(i)' We consider mainly two group comparison.

(ii)' We explicitly state the prior distribution where $\Pr(H_1 \mid x) = 1 - p$ is achieved.

(iii)' We consider Poisson and binomial distribution whose probability functions are asymmetric about 0.

2.2 Bayesian probabilities being the hypothesis true for the specific distributions

Next, consider the posterior probabilities $\Pr(H_i \mid x) \quad (i = 0, 1)$ for more concrete situations, that is, for the superiority and non-inferiority of the parameters of normal, binomial, and Poisson distributions.

2.2.1 Mean of the normal distribution

2.2.1.1 One group

Following Berger (2013), we consider the evaluaion of the mean of the normal distribution. Let $X_1, \ldots, X_n \overset{iid}{\sim} N(\theta, \sigma^2)$ with σ^2 known and the prior of θ be $\pi(\theta) \propto 1$. Let the observed data of X_i be x_i for $i = 1, \ldots, n$ and $\bar{x} = 1/n \sum_{i=1}^{n} x_i$. Then, the posterior distribution of θ is $N(\bar{x}, \sigma^2/n)$. Here, for the fixed value θ_0, posterior probability of $\theta > \theta_0$ being true is the following:

$$\begin{aligned}
\Pr(\theta > \theta_0 \mid \bar{x}) &= \int_{\theta_0}^{\infty} \sqrt{\frac{n}{2\sigma^2\pi}} \exp\left(-\frac{n(\theta - \bar{x})^2}{2\sigma^2}\right) d\theta \\
&= \int_{\sqrt{n}(\theta_0 - \bar{x})/\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\tilde{\theta}^2}{2}\right) d\tilde{\theta} \qquad \left(\because \tilde{\theta} = \frac{\sqrt{n}(\theta - \bar{x})}{\sigma}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\sqrt{n}(\theta_0-\bar{x})/\sigma}^{\infty} \phi(\tilde{\theta}) d\tilde{\theta} \\
&= 1 - \int_{-\infty}^{\sqrt{n}(\theta_0-\bar{x})/\sigma} \phi(\tilde{\theta}) d\tilde{\theta} \\
&= 1 - \Phi\left(\frac{\sqrt{n}(\theta_0 - \bar{x})}{\sigma}\right) \qquad (2.1)
\end{aligned}
$$

where $\phi(x)$ and $\Phi(x)$ are the probability density function and the cumulative distribution function of the standard normal distribution, respectively. On the other hand, for the null hypothesis $H_0 : \theta \leq \theta_0$ and the alternative $H_1 : \theta > \theta_0$, the p-value based on $Z = \sqrt{n}(\bar{X} - \theta_0)/\sigma \sim N(0,1)$ is

$$
\begin{aligned}
p &= \Pr\left(Z \geq \frac{\sqrt{n}(\bar{x} - \theta_0)}{\sigma}\right) \\
&= 1 - \Phi\left(\frac{\sqrt{n}(\bar{x} - \theta_0)}{\sigma}\right) \\
&= \Phi\left(\frac{\sqrt{n}(\theta_0 - \bar{x})}{\sigma}\right). \qquad (2.2)
\end{aligned}
$$

Then, from (2.1) and (2.2), we obtain

$$
\Pr(H_1 \mid \bar{x}) = 1 - p.
$$

2.2.1.2 Two group comparison (superiority)

Let $X_{ij} \overset{ind}{\sim} N(\mu_i, \sigma^2)$ with σ^2 known, and let x_{ij} be the realized value of X_{ij} for $i = 1, 2$ and $j = 1, \ldots, n$. Let $\bar{X}_i = 1/n \sum_{j=1}^{n} X_{ij}$ and $\bar{x}_i = 1/n \sum_{j=1}^{n} x_{ij}$ for $i = 1, 2$. Suppose the prior distributions of μ_1, μ_2 be $f(\mu_1) \propto 1, f(\mu_2) \propto 1$, then the posterior distributions are $N(\bar{x}_1, \sigma^2/n)$ and $N(\bar{x}_2, \sigma^2/n)$, respectively. Then, $\mu_1 - \mu_2 \mid \bar{x}_1, \bar{x}_2 \sim N(\bar{x}_1 - \bar{x}_2, 2\sigma^2/n)$. Therefore, the posterior probability of superiority hypothesis $\mu_1 > \mu_2$ being true is the following:

$$
\begin{aligned}
\Pr(\mu_1 > \mu_2 \mid \bar{x}_1, \bar{x}_2) &= \Pr(\mu_1 - \mu_2 > 0 \mid \bar{x}_1, \bar{x}_2) \\
&= \int_0^{\infty} \sqrt{\frac{n}{4\pi\sigma^2}} \exp\left(-\frac{n\{z - (\bar{x}_1 - \bar{x}_2)\}^2}{4\sigma^2}\right) dz \qquad (\because z := \mu_1 - \mu_2) \\
&= \int_{-\sqrt{n/2\sigma^2}(\bar{x}_1 - \bar{x}_2)}^{\infty} \phi(z_1) dz_1 \qquad \left(\because z_1 := \sqrt{\frac{n}{2\sigma^2}}\{z - (\bar{x}_1 - \bar{x}_2)\}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \int_{-\infty}^{\sqrt{n/2\sigma^2}(\bar{x}_1 - \bar{x}_2)} \phi(z_1) dz_1 \quad (\because z_1 = -z_2) \\
&= \Phi\left(\frac{\sqrt{n}(\bar{x}_1 - \bar{x}_2)}{\sqrt{2}\sigma}\right). \qquad (2.3)
\end{aligned}
$$

On the other hand, when the null and alternative hypotheses are $H_0 : \mu_1 \leq \mu_2$ and $H_1 : \mu_1 > \mu_2$, the p-value based on $Z = \sqrt{n}(\bar{X}_1 - \bar{X}_2)/\sqrt{2}\sigma \sim N(0, 1)$ is

$$
\begin{aligned}
p &= \Pr\left(Z \geq \frac{\sqrt{n}(\bar{x}_1 - \bar{x}_2)}{\sqrt{2}\sigma}\right) \\
&= 1 - \Phi\left(\frac{\sqrt{n}(\bar{x}_1 - \bar{x}_2)}{\sqrt{2}\sigma}\right). \qquad (2.4)
\end{aligned}
$$

From (2.3) and (2.4), we obtain

$$
\Pr(H_1 \mid \bar{x}_1, \bar{x}_2) = 1 - p.
$$

2.2.1.3 Two group comparison (non-inferiority)

Next, consider the non-inferiority of the mean of one normal distribution to the other. Let $X_{ij} \overset{ind}{\sim} N(\mu_i, \sigma_i^2)$ for $i = 1, 2$ and $j = 1, \ldots, n_i$, and $\bar{X}_i = 1/n_i \sum_{j=1}^{n_i} X_{ij}$. Gamalo et al. (2016) evaluated $P(\mu_1 - \mu_2 \geq -\Delta \mid \bar{X}_1, \bar{X}_1)$.

Suppose σ_1^2 and σ_2^2 are known, the posterior distribution of μ_i given $\bar{X}_i$ is $N(\widetilde{\mu}_i, \widetilde{\sigma}_i)$ for $i = 1, 2$, and $\mu_1 \mid \bar{X}_1$ and $\mu_2 \mid \bar{X}_2$ are independent. Then,

$$
\Pr(\mu_1 - \mu_2 \geq -\Delta \mid \bar{X}_1, \bar{X}_2) = 1 - \Phi\left(\frac{-\Delta - (\widetilde{\mu}_1 - \widetilde{\mu}_2)}{\sqrt{\widetilde{\sigma}_1^2 + \widetilde{\sigma}_2^2}}\right).
$$

When the variances σ_1^2 and σ_2^2 are unknown, Gamalo et al. (2016) utilized the Monte Carlo approximation assuming that the prior distributions of σ_1 and σ_2 are inverse gamma distributions.

2.2.1.4 Three group comparison (non-inferiority)

Next, consider the situation where μ_1 is the mean of the new drug group, μ_2 and μ_3 are those of the active control drug group, and the placebo group, respectively. Ghosh et al. (2011) evaluated

$$\Pr\left(\frac{\mu_1-\mu_3}{\mu_2-\mu_3}>\theta \,\middle|\, \mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3\right) \tag{2.5}$$

where θ is the threshold, and $\mathbf{X}_i = (X_{i1},\ldots,X_{in_i})'$ is the data for each group for $i = 1,2,3$. Suppose $X_{ij} \overset{ind}{\sim} N(\mu_i,\sigma_i^2)$ and, for the prior, suppose $\mu_i \mid \sigma_i^2 \sim N(\mu_{0i},\sigma_i^2/\kappa_{0i})$ and $\sigma_i^2 \sim Inv\text{-}Ga(\nu_{0i}/2,\sigma_{0i}^2\nu_{0i}/2)$ conditioned on $\mu_2-\mu_3 > 0$ with fixed parameters $\mu_{0i},\kappa_{0i},\nu_{0i},\sigma_{0i}$ for $i = 1,2,3$ and $j = 1,\ldots,n_i$, which indicates that the active control drug is superior to the placebo group. For the evaluation of (2.5), Ghosh et al. (2011) utilized the Monte Carlo approximation. Next, they considered the case where X_{ij} does not follow the normal distribution utilizing the Dirichlet Process Mixture.

Next, consider to compare (i) μ_1 and μ_2, and (ii) μ_2 and μ_3, simultaneously. Here, we evaluate simultaneously (a) the non-inferiority of the new drug to the active control drug and (b) the superiority of the active control drug to placebo (called assay sensitivity). In this situation, Ghosh et al. (2016) evaluated

$$\Pr(\mu_1-\mu_2 \geq \Delta \cap \mu_2-\mu_3 \geq \Delta/r \mid \mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3)$$

based on the bivariate normal distribution. This probability is calculated by using Monte Carlo approximation.

2.2.2 Binomial probability

2.2.2.1 One group

Following Zaslavsky (2010), we first consider the one group situation. Let X follow the binomial distribution $Bin(n,\pi)$ and let the prior distribution of π be $Beta(\alpha,\beta)$ for $\alpha,\beta > 0$. Then, given the data $X = x$, the posterior distribution of π is $Beta(a,b)$ where $a = \alpha + x, b = \beta + (n-x)$. If

$a, b \in \mathbb{N}$,

$$\begin{aligned} \Pr(\pi \geq p \mid x) &= \frac{1}{B(a,b)} \int_p^1 \pi^{a-1}(1-\pi)^{b-1} d\pi \\ &= \sum_{r=0}^{a-1} \binom{a+b-1}{r} p^r (1-p)^{a+b-1-r}. \end{aligned} \tag{2.6}$$

For the frequentist hypothesis testing with $H_0 : \pi \geq p$ and $H_1 : \pi < p$, the p-value is given as follows:

$$P_F(x) = \Pr(X \leq x \mid \pi = p) = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}. \tag{2.7}$$

From (2.6) and (2.7), when $x = a - 1$ and $n = a + b - 1$, that is, the posterior distribution of π is $Beta(x+1, n-x)$, $P_F(x)$ equals $\Pr(\pi \geq p \mid x)$. Remark that, for this situation, the corresponding prior of π is not a beta distribution but the improper prior $f(\pi) \propto (1-\pi)^{-1}$.

2.2.2.2 Two group comparison (superiority)

Next, consider two group comparison. Let $X_i \overset{ind}{\sim} Bin(n_i, \pi_i)$ and let the prior distribution of π_i be $Beta(\alpha_i, \beta_i)$ for $i = 1, 2$ and let π_1 and π_2 be independent. Given the data $X_i = x_i$, the posterior distribution of π_i is $Beta(a_i, b_i)$, where $a_i = \alpha_i + x_i, b_i = \beta_i + (n_i - x_i)$ for $i = 1, 2$. Here, to evaluate $\Pr(\pi_1 > \pi_2 \mid X_1, X_2)$, Kawasaki and Miyaoka (2012b) utilized the following normal approximation.

$$\Pr(\pi_1 > \pi_2 \mid X_1, X_2) \approx 1 - \Phi\left(\frac{-\left(\frac{a_1}{a_1+b_1} - \frac{a_2}{a_2+b_2}\right)}{\sqrt{\frac{a_1 b_1}{(a_1+b_1)^2(a_1+b_1+1)} + \frac{a_2 b_2}{(a_2+b_2)^2(a_2+b_2+1)}}} \right).$$

Kawasaki and Miyaoka (2012b) also derived the following expression for the posterior probability:

$$\Pr(\pi_1 > \pi_2 \mid X_1, X_2) = \frac{B(a_1+a_2, b_1)}{a_2 B(a_1, b_1) B(a_2, b_2)} \cdot {}_3F_2(a_2, 1-b_2, a_1+a_2; 1+a_2, a_1+a_2+b_1; 1)$$

where

$$ {}_3F_2(k_1,k_2,k_3;l_1,l_2;z)=\sum_{t=0}^{\infty}\frac{(k_1)_t(k_2)_t(k_3)_t}{(l_1)_t(l_2)_t}\cdot\frac{z^t}{t!} $$

is the generalized hypergeometric series.

Next, following theorem states the relationship between the Bayesian posterior probability and the p-value of Fisher's exact test.

Theorem 2.3 (Kawasaki et al. (2014), Altham (1969), Howard (1998))**.** Suppose that the priors of π_1 and π_2 are $f(\pi_1)\propto\pi_1^{-1}$ and $f(\pi_2)\propto(1-\pi_2)^{-1}$, respectively, and $X_1>0, X_2<n_2$, then between the Bayesian posterior probability Pr$(\pi_1>\pi_2\mid X_1,X_2)$ and the one sided p-value of Fisher's exact test with $H_0:\pi_1\le\pi_2$ versus $H_1:\pi_1>\pi_2$, the following relation holds

$$ \text{Pr}(\pi_1>\pi_2\mid X_1,X_2)=1-p. $$

2.2.2.3 Two group comparison (non-inferiority)

Next, consider the two group non-inferiority situation. Let $X_i\overset{ind}{\sim}Bin(n_i,\pi_i)$ for $i=1,2$ and let the priors of π_1 and π_2 be beta distributions and be independent. Gamalo et al. (2011) evaluated Pr$(\pi_1-\pi_2>-\Delta\mid X_1,X_2)$ using Monte Carlo approximation and two type of normal approximations. For the first approximation, they utilized that $Beta(\alpha,\beta)$ is approximated by $N\left(\alpha/(\alpha+\beta),\alpha\beta/\{(\alpha+\beta)^2(\alpha+\beta+1)\}\right)$. For the second one, they approximated the posterior of π_i given $X_i=x_i$ by $N(x_i/n_i,x_i(n_i-x_i)/n_i^3)$ since $\widehat{\pi}_i=X/n\,\dot{\sim}\,N(\pi_i,\pi_i(1-\pi_i)/n_i)$.

When the posterior distirbution of π_i is $Beta(a_i,b_i)$ for $i=1,2$, Kawasaki and Miyaoka (2010) derived the following probability density function of the posterior distribution of $\delta:=\pi_1-\pi_2$:

$$ f(\delta)=\begin{cases}\dfrac{B(a_1,b_2)(1+\delta)^{a_1+b_2-1}}{B(a_1,b_1)B(a_2,b_2)}F_3(a_1,b_2,1-b_1,1-a_2;a_1+b_2;1+\delta,1+\delta) & (-1\le\delta<0)\\ \dfrac{B(a_1+a_2-1,b_1+b_2-1)}{B(a_1,b_1)B(a_2,b_2)} & (\delta=0)\\ \dfrac{B(a_2,b_1)(1-\delta)^{a_2+b_1-1}}{B(a_1,b_1)B(a_2,b_2)}F_3(a_2,b_1,1-b_2,1-a_1;a_2+b_1;1-\delta,1-\delta) & (0<\delta\le1)\end{cases} $$

where

$$F_3(k_1, k_2, l_1, l_2; h; u_1, u_2) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{(k_1)_i (k_2)_j (l_1)_i (l_2)_j}{(h)_{i+j}} \cdot \frac{u_1^i}{i!} \cdot \frac{u_2^j}{j!}$$

is the Appell hypergeometric function.

Using $f(\delta)$, Kawasaki and Miyaoka (2013) expressed the posterior probability of non-inferiority hypothesis being true as follows:

$$\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) = \Pr(\delta > -\Delta \mid X_1, X_2) = \int_{-\Delta}^{1} f(\delta) d\delta.$$

This expression is exact. However, it is too complicated. Another expression under some mild assumptions are derived and the relationship between this posterior probability and the p-value of Fisher's exact test are shown in Chapter 6.

On the other hand, Zaslavsky (2013) derived the approximate expression given $X_i = x_i$ and the prior of π_i of being $Beta(a_i, b_i)$ for $i = 1, 2$. When $x_1, x_2, a_1, a_2, b_1, b_2 \in \mathbb{N}$, $x_1 + a_1 > 0, x_2 + a_2 > 0$ and $n_1 + b_1 - x_1 - 1 > 0$,

$$\begin{aligned}
&\Pr(\pi_1 + \Delta \le \pi_2 \mid x_1, x_2) \\
&= \sum_{s=0}^{x_1} \frac{\binom{n_1 + a_1 + b_1 - 1}{s + a_1} \binom{n_2 + a_2 + b_2 - 1}{s + a_2 - 1}}{\binom{n_1 + a_1 + b_1 - 1 + n_2 + a_2 + b_2 - 1}{n_1 + a_1 + b_1 - 1}} \\
&\quad - \Delta \cdot \frac{\Gamma(a_1 + b_1 + n_1)\Gamma(a_2 + b_2 + n_2)}{\Gamma(a_1 + x_1)\Gamma(b_1 + n_1 - x_1)\Gamma(a_2 + x_2)\Gamma(b_2 + n_2 - x_2)} \\
&\qquad \times \frac{\Gamma(a_1 + a_2 + x_1 + x_2 - 1)\Gamma(b_1 + b_2 + n_1 + n_2 - x_1 - x_2 - 1)}{\Gamma(a_1 + a_2 + b_1 + b_2 + n_1 + n_2 - 2)} + o(\Delta).
\end{aligned}$$

As another type of non-inferiority, Zaslavsky (2013) considered $\Pr(\pi_1 \le c\pi_2 \mid X_1, X_2)$ for $0 < c \le 1$ and derived the following exact formula when $r = x_1 + a_1, n = n_1 + a_1 + b_1 - 1, s = x_2 + a_2$ and $m = n_2 + a_2 + b_2 - 1$:

$$\Pr(\pi_1 \le c\pi_2 \mid X_1, X_2) = \sum_{\mu=r}^{n} \sum_{k=0}^{n-\mu} (-1)^{n-\mu-k} c^{n-k} \cdot \frac{\Gamma(n+1)}{\Gamma(\mu+1)\Gamma(n-\mu+1)} \cdot \frac{\Gamma(n-\mu+1)}{\Gamma(k+1)\Gamma(n-\mu-k+1)}$$

$$\times \frac{\Gamma(m+n+s-k)}{\Gamma(s)\Gamma(m+n-k+1)} \cdot \frac{\Gamma(m+1)\Gamma(n+s-k)}{\Gamma(m+n+s-k)}.$$

2.2.3 Poisson rate parameter

2.2.3.1 One group

Following Zaslavsky (2010), we consider the one group situation. Let X follow the Poisson distribution $Po(n\lambda)$ with $n \in \mathbb{N}$ and $\lambda > 0$. Let the prior distribution of λ be $Ga(\alpha, \beta)$ with $\alpha, \beta > 0$. Given the observed data $X = x$, the posterior distribution of λ is $Ga(a, b)$ where $a = \alpha + x, b = \beta + n$. If $a \in \mathbb{N}$, for the fixed λ_0,

$$\begin{aligned}
\Pr(\lambda < \lambda_0 \mid x) &= \int_0^{\lambda_0} \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp(-b\lambda) d\lambda \\
&= 1 - \int_{\lambda_0}^{\infty} \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp(-b\lambda) d\lambda \\
&= 1 - \int_{b\lambda_0}^{\infty} \frac{1}{\Gamma(a)} \lambda_1^{a-1} \exp(-\lambda_1) d\lambda_1 \quad (\because \lambda_1 = b\lambda) \\
&= \sum_{m=0}^{a-1} \frac{(b\lambda_0)^m}{m!} \exp(-b\lambda_0). \qquad (2.8)
\end{aligned}$$

For the frequentist p-value with $H_0 : \lambda \geq \lambda_0$ vs $H_1 : \lambda < \lambda_0$, given $X = x$, is

$$P_F(x) = \Pr(X \leq x \mid \lambda = \lambda_0) = \sum_{m=0}^{x} \frac{(n\lambda_0)^m}{m!} \exp(-n\lambda_0). \qquad (2.9)$$

If the prior distribution of λ is $Ga(1, 1)$, the posterior distribution of λ is $Ga(1+x, 1+n)$. Therefore,

$$\Pr(\lambda < \lambda_0 \mid x) = \sum_{m=0}^{x} \frac{\{(n+1)\lambda_0\}^m}{m!} \exp(-b\lambda_0).$$

Furthermore, if the posterior distribution of λ is $Ga(1 + x, n)$,

$$\Pr(\lambda < \lambda_0 \mid x) = \sum_{m=0}^{x} \frac{(n\lambda_0)^m}{m!} \exp(-b\lambda_0) = P_F(x).$$

Remark that, in this case, the corresponding prior distribution is not a gamma distribution, but the improper prior $f(\lambda) \propto \lambda^{-1}$.

2.2.3.2 Two group comparison (superiority)

Let $X_i \overset{ind}{\sim} Po(n_i\lambda_i)$ with $n_i, \lambda_i > 0$ and let the prior distribution of λ_i be $Ga(\alpha_i, \beta_i)$ with $\alpha_i, \beta_i > 0$ for $i = 1, 2$. Given $X_i = x_i$, the posterior distribution of λ_i is $Ga(a_i, b_i)$ where $a_i = \alpha_i + x_i$ and $b_i = \beta_i + n_i$. In this situation, Kawasaki and Miyaoka (2012a) derived the following approximation formula for $\Pr(\lambda_1 < \lambda_2 \mid X_1, X_2)$:

$$\Pr(\lambda_1 < \lambda_2 \mid X_1, X_2) \approx \Phi\left(\frac{-a_1/b_1 + a_2/b_2}{\sqrt{a_1/b_1^2 + a_2/b_2^2}}\right).$$

Furthermore, Kawasaki and Miyaoka (2012a) also derived the following “exact” expression:

$$\Pr(\lambda_1 < \lambda_2 \mid X_1, X_2) = 1 - \frac{1}{a_2 B(a_1, a_2)}\left(\frac{b_2}{b_1 + b_2}\right)^{a_2} \cdot {}_2F_1\left(a_2, 1 - a_1; 1 + a_2; \frac{b_2}{b_1 + b_2}\right).$$

For this probability, other quite simple exact expressions and the relation to the frequentist p-value is shown in Chapter 3.

2.2.3.3 Two group comparison (non-inferiority)

Kawasaki et al. (2016) evaluated the posterior probability $\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2)$ by normal approximation

$$\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) \approx \Phi\left(\frac{\Delta - (a_1/b_1 - a_2/b_2)}{\sqrt{a_1/b_1^2 + a_2/b_2^2}}\right)$$

and Monte Carlo approximation. For this probability, exact expression under some mild conditions is derived in Chapter 4 .

2.3 Bayes factor

2.3.1 Definition

Another famous evidence for Bayesian decision making is the Bayes factor. Let the parameter of interest be θ and the parameter space be Θ. Then, let Θ_0 and Θ_1 be the parameter spaces corresponding to the null and alternative hypothesis, that is,

$$H_0 : \theta \in \Theta_0, \quad H_1 : \theta \in \Theta_1,$$

where $\Theta_0 \cap \Theta_1 = \emptyset$. Here, we suppose $\Theta_0 \cup \Theta_1 = \Theta$.

Next, let

$$\Pr(H_i) = \int_{\Theta_i} \pi(\theta) d\theta$$

be the prior probability of H_i being true.

Following Robert (2007), Bayes factor is defined as follows

Definition 2.4. The Bayes factor is the ratio of the posterior probabilities of the null hypothesis H_0 and the alternative hypotheses H_1 over the ratio of the prior probabilities of the null and alternative hypotheses, i.e.,

$$BF_{01}(x) := \frac{\Pr(H_0 \mid x)}{\Pr(H_1 \mid x)} \Big/ \frac{\Pr(H_0)}{\Pr(H_1)}. \tag{2.10}$$

2.3.2 Relation between the Bayes factor and the posterior probability of the hypothesis being true

Since $\Theta = \Theta_0 \cup \Theta_1$, $\Pr(H_1) = 1 - \Pr(H_0)$ and $\Pr(H_1 \mid x) = 1 - \Pr(H_0 \mid x)$. Therefore, from (2.10), we obtain the following relationship between the Bayes factor and the posterior probability of the hypothesis being true:

$$\Pr(H_0 \mid x) = \frac{1}{1 + \dfrac{1 - \Pr(H_0)}{\Pr(H_0)} \cdot \dfrac{1}{BF_{01}(x)}}.$$

For the detail of Bayes factors, see Jeffreys (1961), Kass and Raftery (1995), Berger et al. (2001), Ghosh et al. (2005), Pericchi (2005) and references therein.

Chapter 3

Bayesian superiority and non-inferiority hypothesis testings and the p-value of the conditional test for the Poisson rate parameters

3.1 Introduction

Comparing two groups is one of the most popular topics in statistics. For comparisons, frequentist methods are often applied in medical statistics and epidemiology. However, in recent years, Bayesian methods have gained increasing attention because prior information can be used to improve the efficiency of inference. Particularly for categorical data analysis, to evaluate the superiority of one group over another from the Bayesian perspective, Kawasaki and Miyaoka (2012a,b, 2014), and Howard (1998) investigated the posterior probability $\Pr(\theta_1 > \theta_2 \mid X_1, X_2)$ where X_1, X_2 are data and θ_1, θ_2 are parameters of interest in each group. For binomial distributions with proportions π_1, π_2, Kawasaki and Miyaoka (2012b) referred to $\theta = \Pr(\pi_1 > \pi_2 \mid X_1, X_2)$ as a Bayesian index and expressed it by the hypergeometric series. Kawasaki et al. (2014) showed that θ and the one sided p-value of Fisher's exact test are equivalent under certain conditions. A similar relationship was investigated by Altham (1969) and Howard (1998). For Poisson distributions with parameters λ_1, λ_2, Kawasaki and Miyaoka (2012a) proposed a Bayesian index $\theta = \Pr(\lambda_1 < \lambda_2 \mid X_1, X_2)$, expressed it using the hypergeometric series, and inferred the relationship between θ and the one-sided p-value of the z-type Wald test. However, hypergeometric series are, in general, difficult to calculate, and the exact relationship between θ and p-value was not established. In this chapter, we give other expressions for the Bayesian index, which can be easily calculated, and show the exact relationship between θ with the non-informative prior and the one-sided p-value of the conditional test. Additionally, we investigate the relationship between the generalized version of the Bayesian index and the p-value of the conditional test with more general hypotheses. The remainder of this chapter is structured as follows. In Section 3.2, we give four expressions for the Bayesian index $\Pr(\lambda_1 < \lambda_2 \mid X_1, X_2)$ other than the hypergeometric series under some conditions. In Section 3.3, we investigate the relationship between the Bayesian index and the p-value of the conditional test with the null hypothesis $H_0 : \lambda_1 \geq \lambda_2$ versus the alternative hypothesis $H_1 : \lambda_1 < \lambda_2$. In Section 3.4, as a generalization, we investigate the relationship between $\Pr(\lambda_1/\lambda_2 < c)$ and the p-value of the conditional test with the null hypothesis $H_0 : \lambda_1/\lambda_2 \geq c$ versus the alternative $H_1 : \lambda_1/\lambda_2 < c$. In Section 3.6, we illustrate the Bayesian index using analyses of real data. Finally, we provide some concluding remarks in Section 3.7.

3.2 Bayesian index for the Poisson parameters and its expressions

3.2.1 Bayesian index with the gamma prior

We consider two situations. First, for $i = 1, 2$ and $j = 1, \ldots, n_i$, let X_{ij} be the outcome of jth subject in the ith group and independently follow the Poisson distribution $Po(\lambda_i)$, and let $X_i = \sum_{j=1}^{n_i} X_{ij}$. Second, for $i = 1, 2$, let X_i be the independent Poisson process with Poisson rate λ_i and let n_i be the person-years at risk. For both cases, $X_i \overset{ind}{\sim} Po(n_i\lambda_i)$. In the following, let n_1, n_2 be the fixed integers for simplicity. For the Bayesian analysis, let the prior distributions of λ_i be $Ga(\alpha_i, \beta_i)$ with $\alpha_i, \beta_i > 0$ for $i = 1, 2$, whose probability density function is $f(\lambda_i \mid \alpha_i, \beta_i) = \beta_i^{\alpha_i}/\Gamma(\alpha_i) \cdot \lambda_i^{\alpha_i - 1} \exp(-\beta_i\lambda_i)$. Let $X_i = k_i$, $a_i := \alpha_i + k_i$, and $b_i := \beta_i + n_i$, then the posterior distributions of λ_i is $Ga(a_i, b_i)$. Here, if $\alpha_i \in \mathbb{N}$, then $a_i \in \mathbb{N}$ for $i = 1, 2$. However, in the following, we suppose that $\alpha_1, \alpha_2, \beta_1, \beta_2 > 0$ and $k_1, k_2 \in \mathbb{N} \cup \{0\}$. When the posterior of λ_i is $Ga(a_i, b_i)$ for $i = 1, 2$, Kawasaki and Miyaoka (2012a) proposed the Bayesian index $\theta = \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ and derived the following expression:

$$\begin{aligned} & \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) \\ = \;& 1 - \frac{1}{a_2 B(a_1, a_2)} \left(\frac{b_2}{b_1 + b_2} \right)^{a_2} \cdot {}_2F_1 \left(a_2, 1 - a_1; 1 + a_2; \frac{b_2}{b_1 + b_2} \right) \end{aligned} \tag{3.1}$$

where

$${}_2F_1(a, b; c; z) = \sum_{t=0}^{\infty} \frac{(a)_t (b)_t}{(c)_t} \cdot \frac{z^t}{t!} \qquad (|z| < 1)$$

is the hypergeometric series and $(k)_t$ is the Pochhammer symbol, that is, $(k)_0 = 1$ and $(k)_t = k(k+1) \cdot (k+t-1)$ for $t \in \mathbb{N}$. Let $F_{\nu_1, \nu_2}(x)$ be the cumulative distribution function of F distribution with degrees of freedom (ν_1, ν_2), that is,

$$F_{\nu_1, \nu_2}(x) = \int_0^x \frac{1}{zB\left(\nu_1/2, \nu_2/2\right)} \left(\frac{\nu_1 z}{\nu_1 z + \nu_2} \right)^{\nu_1/2} \left(\frac{\nu_2}{\nu_1 z + \nu_2} \right)^{\nu_2/2} dz \tag{3.2}$$

where $B(a, b) = \int_0^\infty x^{a-1}(1-x)^{b-1} dx$ is the beta function. Then, we can obtain the following expressions for θ.

Theorem 3.1. If the posterior distribution of λ_i is $Ga(a_i, b_i)$ with $a_i, b_i > 0$ for $i = 1, 2$, then the Bayesian index $\theta = \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ has the following two expressions:

$$\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) = I_{b_1/(b_1+b_2)}(a_1, a_2) \tag{3.3}$$

$$= F_{2a_1, 2a_2} \left(\frac{b_1/a_1}{b_2/a_2} \right) \tag{3.4}$$

where

$$I_x(a, b) = \frac{1}{B(a, b)} \int_0^x t^{a-1}(1-t)^{b-1} dt$$

is the cumulative distribution function of the beta distribution, also known as the regularized incomplete beta function. Moreover, if both $a_1$and a_2 are natural numbers, then $\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$

has the following two additional expressions:

$$\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) = \sum_{r=0}^{a_2-1} \binom{a_1+a_2-1}{r} \left(\frac{b_2}{b_1+b_2}\right)^r \left(\frac{b_1}{b_1+b_2}\right)^{a_1+a_2-1-r} \tag{3.5}$$

$$= \sum_{r=0}^{a_2-1} \binom{a_1+r-1}{a_1-1} \left(\frac{b_1}{b_1+b_2}\right)^{a_1} \left(\frac{b_2}{b_1+b_2}\right)^r. \tag{3.6}$$

(3.5) and (3.6) are the cumulative distribution functions of the binomial and negative binomial distributions, respectively.

Proof. First, (3.3) can be shown by modifying (3.1) using $I_x(a,b) = \frac{1}{B(a,b)} \cdot \frac{x^a}{a} \cdot {}_2F_1(a, 1-b; 1+a; x)$ and $I_z(a,b) = 1 - I_{1-z}(b,a)$, which are 26.5.23 and 26.5.2 of Abramowitz and Stegun (1964), respectively. Next, (3.4) can be shown by changing variable $t = \nu_1 z/(\nu_1 z + \nu_2)$ for (3.2) with $\nu_1 = 2a_1, \nu_2 = 2a_2$ and $x = (b_1/a_1)/(b_2/a_2)$. In the following, suppose that $a_1, a_2 \in \mathbb{N}$. (3.5) can be shown by (3.3), 26.5.2 of Abramowitz and Stegun (1964) above, and 26.5.4 of Abramowitz and Stegun (1964): $\sum_{r=a}^{n} \binom{n}{r} p^r (1-p)^{n-r} = I_p(a, n-a+1)$ as follows

$$\begin{aligned}
\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) &= I_{b_1/(b_1+b_2)}(a_1, a_2) \quad (\because (3.3)) \\
&= 1 - I_{b_2/(b_1+b_2)}(a_2, a_1) \quad (\because 26.5.2 \text{ of Abramowitz and Stegun (1964)}) \\
&= 1 - \sum_{r=a_2}^{a_1+a_2-1} \binom{a_1+a_2-1}{r} \left(\frac{b_2}{b_1+b_2}\right)^r \left(\frac{b_1}{b_1+b_2}\right)^{a_1+a_2-1-r} \\
&\qquad (\because \text{26.5.4 of Abramowitz and Stegun (1964)}) \\
&= \sum_{r=0}^{a_2-1} \binom{a_1+a_2-1}{r} \left(\frac{b_2}{b_1+b_2}\right)^r \left(\frac{b_1}{b_1+b_2}\right)^{a_1+a_2-1-r}.
\end{aligned}$$

Finally, (3.6) can be shown by 8.352-2 of Zwillinger (2014) :

$$\int_x^\infty e^{-t} t^n dt = n! \cdot e^{-x} \sum_{m=0}^{n} \frac{x^m}{m!} \quad (n \in \mathbb{N}, x \in \mathbb{R}) \tag{3.7}$$

as follows

$$\begin{aligned}
&\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) \\
= &\int_0^\infty \left(\int_{\lambda_1}^\infty \frac{b_2^{a_2}}{\Gamma(a_2)} \lambda_2^{a_2-1} \exp(-b_2\lambda_2) d\lambda_2 \right) \frac{b_1^{a_1}}{\Gamma(a_1)} \lambda_1^{a_1-1} \exp(-b_1\lambda_1) d\lambda_1.
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^\infty \left(\int_{b_2\lambda_1}^\infty \frac{b_2^{a_2}}{\Gamma(a_2)} \left(\frac{\pi_2}{b_2}\right)^{a_2-1} \exp(-\pi_2) \frac{d\pi_2}{b_2} \right) \frac{b_1^{a_1}}{\Gamma(a_1)} \lambda_1^{a_1-1} \exp(-b_1\lambda_1) d\lambda_1 \\
&\qquad\qquad (\because \pi_2 = b_2\lambda_2) \\
&= \int_0^\infty \left(\sum_{r=0}^{a_2-1} \frac{(b_2\lambda_1)^r}{r!} \exp(-b_2\lambda_1) \right) \frac{b_1^{a_1}}{\Gamma(a_1)} \lambda_1^{a_1-1} \exp(-b_1\lambda_1) d\lambda_1 \quad (\because (3.7)) \\
&= \sum_{r=0}^{a_2-1} \frac{b_1^{a_1} b_2^r}{\Gamma(a_1) r!} \int_0^\infty \lambda_1^{a_1+r-1} \exp(-(b_1+b_2)\lambda_1) d\lambda_1 \\
&= \sum_{r=0}^{a_2-1} \frac{b_1^{a_1} b_2^r}{\Gamma(a_1) r!} \int_0^\infty \left(\frac{\pi_1}{b_1+b_2}\right)^{a_1+r-1} \exp(-\pi_1) \cdot \frac{d\pi_1}{b_1+b_2} \quad (\because \pi_1 = (b_1+b_2)\lambda_1) \\
&= \sum_{r=0}^{a_2-1} \frac{\Gamma(a_1+r)}{\Gamma(a_1) r!} \left(\frac{b_1}{b_1+b_2}\right)^{a_1} \left(\frac{b_2}{b_1+b_2}\right)^r \\
&= \sum_{r=0}^{a_2-1} \binom{a_1+r-1}{a_1-1} \left(\frac{b_1}{b_1+b_2}\right)^{a_1} \left(\frac{b_2}{b_1+b_2}\right)^r .
\end{aligned}
$$

We have completed the proof of Theorem 3.1. □

Kawasaki and Miyaoka (2012a) expressed θ using the hypergeometric series and computed it by summing the series. However, in general, it is difficult to calculate the hypergeometric series. Additionally, it is also difficult to understand the relationship between θ and other distributions. On the other hand, our expressions above have two advantages. First, we can calculate θ easily using the cumulative distribution functions of well-known distributions. Second, we can find the relationship between θ and some values represented by these cumulative distribution functions. Particularly, from expression (3.3), we can easily show the relationship between θ and the p-value of the conditional test in Section 3.3. Here, we note an assumption for Theorem 3.1. At the beginning of this section, we supposed the priors to be gamma distributions. However, for Theorem 3.1, we only need the posteriors to be gamma. Therefore, as long as the posteriors are gamma, we need not assume the priors to be gamma.

3.2.2 Examples of the prior distribution

In this section, we consider several examples of the prior of λ_i. All of the following examples are gamma distribution or the limit of the gamma distribution, and all the posteriors are gamma.

Example 1 (Non-informative prior)**.**

The non-informative prior distribution is $f_1(\lambda_i) \propto \lambda_i^{-1}$. This is an improper prior but can be considered the limit of $Ga(\alpha_i, \beta_i)$ when $(\alpha_i, \beta_i) \to (0, 0)$. Here, for $k_i > 0$, the posterior is $Ga(k_i, n_i)$. Therefore, when $k_1, k_2 > 0$, that is, $k_1, k_2 \in \mathbb{N}$, Theorem 3.1 states

$$\begin{aligned}
\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) &= I_{n_1/(n_1+n_2)}(k_1, k_2) \\
&= F_{2k_1,2k_2}\left(\frac{n_1/k_1}{n_2/k_2}\right) \\
&= \sum_{r=0}^{k_2-1} \binom{k_1+k_2-1}{r} \left(\frac{n_2}{n_1+n_2}\right)^r \left(\frac{n_1}{n_1+n_2}\right)^{k_1+k_2-1-r} \\
&= \sum_{r=0}^{k_2-1} \binom{k_1+r-1}{k_1-1} \left(\frac{n_1}{n_1+n_2}\right)^{k_1} \left(\frac{n_2}{n_1+n_2}\right)^{r}.
\end{aligned}$$

When $k_i = 0$, the probability density function of the posterior is

$$f(\lambda_i \,|\, X_i) \;\propto\; \lambda_i^{-1} \cdot \frac{\lambda_i^0}{0!} \exp(-\lambda_i) = \lambda_i^{-1} \exp(-\lambda_i).$$

Hence, the posterior is improper and not a gamma distribution. Therefore, Theorem 3.1 cannot be applied.

Example 2 (Jeffreys prior)**.**

The Jeffreys prior distribution (Jeffreys, 1946) is $f_2(\lambda_i) \propto \lambda_i^{-1/2}$. This is also an improper prior but can be considered the limit of $Ga(1/2, \beta_i)$ when $\beta_i \to 0$. Here, the posterior is $Ga(k_i + 1/2, n_i)$ for $k_i \geq 0$. Therefore, Theorem 3.1 states

$$\begin{aligned}
\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) &= I_{n_1/(n_1+n_2)}\,(k_1 + 1/2, k_2 + 1/2) \\
&= F_{2k_1+1,2k_2+1}\left(\frac{n_1/(k_1+1/2)}{n_2/(k_2+1/2)}\right).
\end{aligned}$$

Because $k_i + 1/2 \notin \mathbb{N}$ for any $k_i \in \mathbb{N} \cup \{0\}$, we cannot have expression (3.5) and (3.6).

Example 3 (Conditional power prior)**.**

Let the historical data $x_i^0 \sim Po(m_i\lambda_i)$. Then the likelihood for λ_i is

$$L(\lambda_i \,|\, x_i^0) = \frac{(m_i\lambda_i)^{x_i^0}}{x_i^0!} \exp(-m_i\lambda_i).$$

Here, an example of the conditional power prior distribution Ibrahim and Chen (2000) is given as

$$f_3(\lambda_i) \propto L(\lambda_i \,|\, x_i^0)^{a_i} \cdot f_1(\lambda_i)$$

where a_i is the fixed parameter such that $0 < a_i \leq 1$ and $f_1(\lambda_i) \propto \lambda_i^{-1}$. Thus

$$\begin{aligned} f_3(\lambda_i) &\propto \left(\lambda_i^{x_i^0} \exp(-m_i\lambda_i)\right)^{a_i} \cdot \lambda_i^{-1} \\ &= \lambda_i^{a_i x_i^0 - 1} \exp(-a_i m_i \lambda_i). \end{aligned}$$

Hence, the prior of λ_i is $Ga(a_i x_i^0, a_i m_i)$ when $x_i^0 > 0$. The posterior is $Ga(a_i x_i^0 + k_i, a_i m_i + n_i)$. Therefore, Theorem 3.1 states

$$\begin{aligned} \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) &= I_{\frac{a_1 m_1 + n_1}{(a_1 m_1 + n_1) + (a_2 m_2 + n_2)}}(a_1 x_1^0 + k_1, a_2 x_2^0 + k_2) \\ &= F_{2(a_1 x_1^0 + k_1), 2(a_2 x_2^0 + k_2)}\left(\frac{(a_1 m_1 + n_1)/(a_1 x_1^0 + k_1)}{(a_2 m_2 + n_2)/(a_2 x_2^0 + k_2)}\right). \end{aligned}$$

Here, because $a_1 x_1^0, a_2 x_2^0 \notin \mathbb{N}$ in general, we cannot have expression (3.5) and (3.6) in general. On the other hand, when $a_1 x_1^0, a_2 x_2^0 \in \mathbb{N}$, we have expression (3.5) and (3.6) as follows

$$\begin{aligned} &\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) \\ = &\sum_{r=0}^{a_2 x_2^0 + k_2 - 1} \binom{a_1 x_1^0 + k_1 + a_2 x_2^0 + k_2 - 1}{r} \\ &\quad \times \left(\frac{a_2 m_2 + n_2}{a_1 m_1 + n_1 + a_2 m_2 + n_2}\right)^r \left(\frac{a_1 m_1 + n_1}{a_1 m_1 + n_1 + a_2 m_2 + n_2}\right)^{a_1 x_1^0 + k_1 + a_2 x_2^0 + k_2 - 1 - r} \\ = &\sum_{r=0}^{a_2 x_2^0 + k_2 - 1} \binom{a_1 x_1^0 + k_1 + r - 1}{a_1 x_1^0 + k_1 - 1} \\ &\qquad \times \left(\frac{a_1 m_1 + n_1}{a_1 m_1 + n_1 + a_2 m_2 + n_2}\right)^{a_1 x_1^0 + k_1} \left(\frac{a_2 m_2 + n_2}{a_1 m_1 + n_1 + a_2 m_2 + n_2}\right)^r. \end{aligned}$$

3.3 The relationship between the Bayesian index and the p-value of the conditional test

For binomial proportions, Kawasaki et al. (2014) and Altham (1969) showed the relationship between the Bayesian index and the one-sided p-value of Fisher's exact test under certain conditions. For Poisson parameters, a similar relationship holds between the Bayesian index and the one-sided p-value of the conditional test.

3.3.1 Conditional test

From the frequentist perspective, we consider the conditional test based on the conditional distribution of X_1 given $X_1 + X_2 = k_1 + k_2$ (Przyborowski and Wilenski (1940); Krishnamoorthy and Thomson (2004)). The probability function is

$$\begin{aligned} & f(X_1 = k_1 \,|\, X_1 + X_2 = k_1 + k_2) \\ & \qquad = \binom{k_1 + k_2}{k_1} \left(\frac{n_1 \lambda_1}{n_1 \lambda_1 + n_2 \lambda_2} \right)^{k_1} \left(\frac{n_2 \lambda_2}{n_1 \lambda_1 + n_2 \lambda_2} \right)^{(k_1 + k_2) - k_1} . \end{aligned}$$

To test the null hypothesis $H_0 : \lambda_1 \geq \lambda_2$ versus the alternative $H_1 : \lambda_1 < \lambda_2$, the p-value is

$$\begin{aligned} p &= \Pr(X_1 \leq k_1 \,|\, X_1 + X_2 = k_1 + k_2, \lambda_1 = \lambda_2) \\ &= \sum_{r=0}^{k_1} \binom{k_1 + k_2}{r} \left(\frac{n_1}{n_1 + n_2} \right)^{r} \left(\frac{n_2}{n_1 + n_2} \right)^{k_1 + k_2 - r} . \end{aligned} \tag{3.8}$$

Lemma 3.2 (Doi (2016))**.** If $k_2 > 0$, then the one-sided p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$ has the following expressions:

$$\begin{aligned} p &= I_{n_2/(n_1+n_2)}(k_2, k_1 + 1) \\ &= F_{2k_2, 2(k_1+1)} \left(\frac{n_2/k_2}{n_1/(k_1 + 1)} \right) \\ &= \sum_{r=0}^{k_1} \binom{k_2 + r - 1}{k_2 - 1} \left(\frac{n_2}{n_1 + n_2} \right)^{k_2} \left(\frac{n_1}{n_1 + n_2} \right)^{r} \\ &= \frac{1}{k_2 B(k_2, k_1 + 1)} \left(\frac{n_2}{n_1 + n_2} \right)^{k_2} \cdot {}_2F_1 \left(k_2, -k_1; 1 + k_2; \frac{n_2}{n_1 + n_2} \right) . \end{aligned}$$

If $k_2 = 0$, then $p = 1$.

Proof. For $k_2 > 0$, the proof is similar to that of Theorem 3.1 with $\alpha_1, \alpha_2 \in \mathbb{N}$. For $k_2 = 0$, from (3.8),

$$p = \sum_{r=0}^{k_1} \binom{k_1}{r} \left(\frac{n_1}{n_1+n_2}\right)^r \left(\frac{n_2}{n_1+n_2}\right)^{k_1-r} = 1.$$

□

3.3.2 The relationship between the Bayesian index and the p-value of the conditional test

Theorem 3.3 (Doi (2016))**.** (1) If $k_2 > 0$, then between $\theta = \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ given $X_1 = k_1 + 1$, $X_2 = k_2$ and the one-sided p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$ given $X_1 = k_1, X_2 = k_2$, the following relation holds

$$\lim_{(\alpha_1,\alpha_2,\beta_1,\beta_2)\to(0,0,0,0)} \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) = 1 - p.$$

Here n_1 and n_2 are the same for $\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ and p.

(2) Suppose that the prior of λ_i is $Ga(\alpha_i, \beta_i)$ with $\alpha_i, \beta_i \in \mathbb{N}$ for $i = 1, 2$, and let $m_1, m_2 \in \mathbb{N}$. Then, between $\theta = \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ given $X_1 = k_1 - \alpha_1 + 1, X_2 = k_2 - \alpha_2,\ n_1 = m_1 - \beta_1, n_2 = m_2 - \beta_2$, and the one-sided p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$ given $X_1 = k_1, X_2 = k_2,\ n_1 = m_1, n_2 = m_2$, the following relation holds

$$\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) = 1 - p.$$

Proof. (1) From Lemma 3.2, the p-value given $X_1 = k_1, X_2 = k_1$ is

$$p = I_{n_2/(n_1+n_2)}(k_2, k_1 + 1).$$

Therefore, from the relation $I_z(a,b) = 1 - I_{1-z}(b,a)$,

$$1 - p = I_{n_1/(n_1+n_2)}(k_1+1, k_2). \tag{3.9}$$

Given $X_1 = k_1+1, X_2 = k_2$, on the other hand, $a_1 = \alpha_1 + k_1 + 1, a_2 = \alpha_2 + k_2, b_1 = \beta_1 + n_1, b_2 = \beta_2 + n_2$. Therefore, the Bayesian index is

$$\begin{aligned} \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) &= I_{b_1/(b_1+b_2)}(a_1, a_2) \\ &= I_{\frac{\beta_1+n_1}{\beta_1+n_1+\beta_2+n_2}}(\alpha_1 + k_1 + 1, \alpha_2 + k_2). \end{aligned} \tag{3.10}$$

From (3.9) and (3.10),

$$\lim_{(\alpha_1,\alpha_2,\beta_1,\beta_2)\to(0,0,0,0)} \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) = 1 - p$$

holds. We have just completed the proof of (1).

(2) From Lemma 3.2, the p-value given $X_1 = k_1, X_2 = k_2, n_1 = m_1, n_2 = m_2$ is

$$p = I_{m_2/(m_1+m_2)}(k_2, k_1 + 1).$$

Therefore,

$$1 - p = I_{m_1/(m_1+m_2)}(k_1 + 1, k_2). \tag{3.11}$$

On the other hand, given $X_1 = k_1 - \alpha_1 + 1, X_2 = k_2 - \alpha_2, n_1 = m_1 - \beta_1, n_2 = m_2 - \beta_2$, and $Ga(\alpha_i, \beta_i)$ as the prior for λ_i for $i = 1, 2$, the posterior of λ_1 and λ_2 are $Ga(k_1 + 1, m_1)$ and $Ga(k_2, m_2)$, respectively. Then, the Bayesian index is

$$\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) = I_{m_1/(m_1+m_2)}(k_1 + 1, k_2). \tag{3.12}$$

From (3.11) and (3.12),

$$\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) = 1 - p$$

holds. We have completed the proof of (2). $\square$

Remark 1. (i) $\lim\limits_{(\alpha_1,\alpha_2,\beta_1,\beta_2)\to(0,0,0,0)} \theta$ equals the Bayesian index with the non-informative prior when $k_1, k_2 > 0$.

(ii) For the conditional power prior, if both $a_i x_i^0$ and $a_i m_i$ are natural numbers for $i = 1, 2$, then (2) of Theorem 3.3 can be applied.

(iii) For the conditional power prior, if at least one of $a_1 x_1^0$ and $a_2 x_2$ is not an integer, then at least one of $a_1 x_1^0 + k_1$ and $a_2 x_2^0 + k_2$ is not a natural number. Therefore, $\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) = I_{\frac{a_1 m_1 + n_1}{(a_1 m_1 + n_1) + (a_2 m_2 + n_2)}}(a_1 x_1^0 + k_1, a_2 x_2^0 + k_2)$ does not equal $1 - p$. Similarly, for the Jeffreys prior, since $k_1 + 1/2$ and $k_2 + 1/2$ are not natural numbers, $\Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2) = I_{n_1/(n_1+n_2)}(k_1 + 1/2, k_2 + 1/2)$ does not equal $1 - p$.

Corollary 3.4. Suppose the prior distribution of λ_1 and λ_2 are $f(\lambda_1) \propto 1$ and $f(\lambda_2) \propto \lambda_2^{-1}$, respectively. Then, if $X_2 > 0$, between the posterior probability and the one-sided p-value of the conditional test, the following relationship holds

$$\Pr(\lambda_1 < \lambda_2 \mid X_1, X_2) = 1 - p.$$

Proof. The proof directly follows from Theorem 3.3. $\square$

3.4 Generalization

As a generalization of Theorems 3.1 and 3.3, we consider the generalized version of the Bayesian index $\theta = \Pr(\lambda_1/\lambda_2 < c \,|\, X_1, X_2)$, and investigate the relationship between θ and the one-sided p-value of the conditional test with the null hypothesis $H_0 : \lambda_1/\lambda_2 \geq c$ versus the alternative $H_1 : \lambda_1/\lambda_2 < c$. Let $\pi := \lambda_1/\lambda_2$. We consider the posterior of π when the posterior of λ_i is $Ga(a_i, b_i)$ for $i = 1, 2$. First, the joint density function of (λ_1, λ_2) is

$$f(\lambda_1 \,|\, a_1, b_1) \cdot f(\lambda_2 \,|\, a_2, b_2) = \frac{b_1^{a_1} b_2^{a_2}}{\Gamma(a_1)\Gamma(a_2)} \lambda_1^{a_1-1} \lambda_2^{a_2-1} \exp(-(b_1\lambda_1 + b_2\lambda_2)).$$

Next, let $\pi_1 = \lambda_2$. Then, $\lambda_1 = \pi \cdot \pi_1$, $\lambda_2 = \pi_1$. Finally, the probability density function of the posterior distribution of π is

$$
\begin{aligned}
f(\pi) &= \int_0^\infty f(\pi \cdot \pi_1 \,|\, a_1, b_1) \cdot f(\pi_1 \,|\, a_2, b_2) \cdot \left| \begin{matrix} \dfrac{\partial \lambda_1}{\partial \pi} & \dfrac{\partial \lambda_1}{\partial \pi_1} \\ \dfrac{\partial \lambda_2}{\partial \pi} & \dfrac{\partial \lambda_2}{\partial \pi_1} \end{matrix} \right| d\pi_1 \\
&= \int_0^\infty \frac{b_1^{a_1} b_2^{a_2}}{\Gamma(a_1)\Gamma(a_2)} (\pi \cdot \pi_1)^{a_1 - 1} \pi_1^{a_2 - 1} \exp(-(b_1 \pi \cdot \pi_1 + b_2 \pi_1)) \cdot \pi_1 d\pi_1 \\
&= \frac{b_1^{a_1} b_2^{a_2} \pi^{a_1 - 1}}{\Gamma(a_1)\Gamma(a_2)} \int_0^\infty \pi_1^{a_1 + a_2 - 1} \exp(-\pi_1(b_1 \pi + b_2)) d\pi_1 \\
&= \frac{b_1^{a_1} b_2^{a_2} \pi^{a_1 - 1}}{\Gamma(a_1)\Gamma(a_2)} \cdot \frac{\Gamma(a_1 + a_2)}{(b_1 \pi + b_2)^{a_1 + a_2}} \\
&= \frac{1}{\pi B(a_1, a_2)} \left(\frac{b_1 \pi}{b_1 \pi + b_2} \right)^{a_1} \left(\frac{b_2}{b_1 \pi + b_2} \right)^{a_2}.
\end{aligned}
$$

Hence, the cumulative distribution function is

$$
\begin{aligned}
F(x) &= \int_0^x \frac{1}{\pi B(a_1, a_2)} \left(\frac{b_1 \pi}{b_1 \pi + b_2} \right)^{a_1} \left(\frac{b_2}{b_1 \pi + b_2} \right)^{a_2} d\pi \qquad (3.13) \\
&= \frac{B_{b_1 x/(b_1 x + b_2)}(a_1, a_2)}{B(a_1, a_2)} \\
&= I_{b_1 x/(b_1 x + b_2)}(a_1, a_2).
\end{aligned}
$$

From this, we can obtain the expressions for the generalized version of the Bayesian index.

Theorem 3.5 (Doi (2016))**.** If the posterior distribution of λ_i is $Ga(a_i, b_i)$ with $a_i, b_i > 0$ for $i = 1, 2$, then, the Bayesian index $\theta = \text{Pr}(\lambda_1/\lambda_2 < c \,|\, X_1, X_2)$ has the following three expressions:

$$
\begin{aligned}
& \text{Pr}(\lambda_1/\lambda_2 < c \,|\, X_1, X_2) \\
= \; & I_{b_1 c/(b_1 c + b_2)}(a_1, a_2) \qquad (3.14) \\
= \; & F_{2a_1, 2a_2}\left(\frac{b_1 c / a_1}{b_2 / a_2} \right) \\
= \; & 1 - \frac{1}{a_2 B(a_1, a_2)} \left(\frac{b_2}{b_1 c + b_2} \right)^{a_2} \cdot {}_2F_1 \left(a_2, 1 - a_1; 1 + a_2; \frac{b_2}{b_1 c + b_2} \right)
\end{aligned}
$$

Additionally, if both $a_1$and a_2 are natural numbers, then, θ has the following two additional expressions:

$$\begin{aligned}\Pr(\lambda_1/\lambda_2 < c \,|\, X_1, X_2) &= \sum_{r=0}^{a_2-1} \binom{a_1+a_2-1}{r} \left(\frac{b_2}{b_1 c + b_2}\right)^r \left(\frac{b_1 c}{b_1 c + b_2}\right)^{a_1+a_2-1-r} \\ &= \sum_{r=0}^{a_2-1} \binom{a_1+r-1}{a_1-1} \left(\frac{b_1 c}{b_1 c + b_2}\right)^{a_1} \left(\frac{b_2}{b_1 c + b_2}\right)^r .\end{aligned}$$

Proof. (3.14) can be shown as follows

$$\begin{aligned}\Pr(\lambda_1/\lambda_2 < c \,|\, X_1, X_2) &= \Pr\left(\lambda_1/\lambda_2 < c \,|\, X_1, X_2\right) \\ &= \Pr\left(\pi < c \,|\, X_1, X_2\right) \\ &= F(c) \\ &= I_{b_1 c/(b_1 c + b_2)}(a_1, a_2) \quad (\because (3.14)).\end{aligned}$$

The remainder of the proof is almost the same as that of Theorem 3.1. □

On the other hand, the one-sided p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq c$ versus $H_1 : \lambda_1/\lambda_2 < c$ (Przyborowski and Wilenski (1940); Krishnamoorthy and Thomson (2004)) is defined as

$$\begin{aligned}p &= \Pr(X_1 \leq k_1 \,|\, X_1 + X_2 = k_1 + k_2, \lambda_1/\lambda_2 = c) \\ &= \sum_{r=0}^{k_1} \binom{k_1+k_2}{r} \left(\frac{n_1 c}{n_1 c + n_2}\right)^r \left(\frac{n_2}{n_1 c + n_2}\right)^{k_1+k_2-r} .\end{aligned}$$

Then, we can obtain the following lemma.

Lemma 3.6 (Doi (2016))**.** If $k_2 > 0$, then the one-sided p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq c$ vs. $H_1 : \lambda_1/\lambda_2 < c$ has the following expressions:

$$\begin{aligned}p &= I_{n_2/(n_1 c + n_2)}(k_2, k_1 + 1) \\ &= F_{2k_2, 2(k_1+1)}\left(\frac{n_2/k_2}{n_1 c/(k_1+1)}\right) \\ &= \frac{1}{k_2 B(k_2, k_1+1)} \left(\frac{n_2}{n_1 c + n_2}\right)^{k_2} \cdot {}_2F_1\left(k_2, -k_1; 1 + k_2; \frac{n_2}{n_1 c + n_2}\right).\end{aligned}$$

If $k_2 = 0$, then $p = 1$.

Proof. The proof is almost the same as Lemma 3.2. □

Finally, we can obtain the generalization of Theorem 3.3.

Theorem 3.7 (Doi (2016))**.** (1) If $k_2 > 0$, then between $\theta = \Pr(\lambda_1/\lambda_2 < c \,|\, X_1, X_2)$ given $X_1 = k_1 + 1$, $X_2 = k_2$ and the one-sided p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq c$ vs. $H_1 : \lambda_1/\lambda_2 < c$ given $X_1 = k_1, X_2 = k_2$, the following relation holds

$$\lim_{(\alpha_1,\alpha_2,\beta_1,\beta_2)\to(0,0,0,0)} \Pr(\lambda_1/\lambda_2 < c \,|\, X_1, X_2) = 1 - p.$$

Here n_1 and n_2 are the same for θ and p.

(2) Suppose that the prior of λ_i is $Ga(\alpha_i, \beta_i)$ with $\alpha_i, \beta_i \in \mathbb{N}$ for $i = 1, 2$, and let $m_1, m_2 \in \mathbb{N}$. Then, between $\theta = \Pr(\lambda_1/\lambda_2 < c \,|\, X_1, X_2)$ given $X_1 = k_1 - \alpha_1 + 1, X_2 = k_2 - \alpha_2,\ n_1 = m_1 - \beta_1, n_2 = m_2 - \beta_2$, and the one-sided p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq c$ vs. $H_1 : \lambda_1/\lambda_2 < c$ given $X_1 = k_1, X_2 = k_2,\ n_1 = m_1, n_2 = m_2$, the following relation holds

$$\Pr(\lambda_1/\lambda_2 < c \,|\, X_1, X_2) = 1 - p.$$

Proof. (1) From Lemma 3.6, the p-value given $X_1 = k_1, X_2 = k_2$ is

$$p = I_{n_2/(n_1 c + n_2)}(k_2, k_1 + 1).$$

Therefore, from the relation $I_z(a, b) = 1 - I_{1-z}(b, a)$,

$$1 - p = I_{n_1 c/(n_1 c + n_2)}(k_1 + 1, k_2). \tag{3.15}$$

On the other hand, from (3.14), the Bayesian index given $X_1 = k_1 + 1, X_2 = k_2$ can be expressed as

$$\begin{aligned} \Pr(\lambda_1/\lambda_2 < c \,|\, X_1, X_2) &= I_{b_1 c/(b_1 c + b_2)}(a_1, a_2) \\ &= I_{\frac{(\beta_1+n_1)c}{(\beta_1+n_1)c+(\beta_2+n_2)}}(\alpha_1 + k_1 + 1, \alpha_2 + k_2). \end{aligned} \tag{3.16}$$

$$(\because a_1 = \alpha_1 + k_1 + 1, a_2 = \alpha_2 + k_2, b_1 = \beta_1 + n_1, b_2 = \beta_2 + n_2)$$

Finally, from (3.16) and (3.15),

$$\lim_{(\alpha_1,\alpha_2,\beta_1,\beta_2)\to(0,0,0,0)} \Pr(\lambda_1/\lambda_2 < c \,|\, X_1, X_2) = 1 - p.$$

holds. We have completed the proof of (1).

(2) From Lemma 3.6, the p-value given $X_1 = k_1, X_2 = k_2, n_1 = m_1, n_2 = m_2$ is

$$p = I_{m_2/(m_1c+m_2)}(k_2, k_1 + 1).$$

Therefore,

$$1 - p = I_{m_1c/(m_1c+m_2)}(k_1 + 1, k_2). \tag{3.17}$$

On the other hand, given $X_1 = k_1 - \alpha_1 + 1, X_2 = k_2 - \alpha_2, n_1 = m_1 - \beta_1, n_2 = m_2 - \beta_2$, and $Ga(\alpha_i, \beta_i)$ as the prior for λ_i for $i = 1, 2$, the posterior of λ_1 and λ_2 are $Ga(k_1 + 1, m_1)$ and $Ga(k_2, m_2)$, respectively. Then, the Bayesian index is

$$\Pr(\lambda_1/\lambda_2 < c \,|\, X_1, X_2) = I_{m_1c/(m_1c+m_2)}(k_1 + 1, k_2) \tag{3.18}$$

From (3.17) and (3.18),

$$\Pr(\lambda_1/\lambda_2 < c \,|\, X_1, X_2) = 1 - p$$

holds. We have completed the proof of (2). □

Remark 2. (i) For the conditional power prior, if both $a_i x_i^0$ and $a_i m_i$ are natural numbers for $i = 1, 2$, then (2) of Theorem 3.7 can be applied.

(ii) For the conditional power prior, if at least one of $a_1 x_1^0$ and $a_2 x_2$ is not an integer, then at least one of $a_1 x_1^0 + k_1$ and $a_2 x_2^0 + k_2$ is not a natural number. Therefore, $\theta = I_{\frac{(a_1 m_1 + n_1)c}{(a_1 m_1 + n_1)c + (a_2 m_2 + n_2)}}(a_1 x_1^0 + k_1, a_2 x_2^0 + k_2)$ does not equal $1 - p$. Similarly, for the Jeffreys prior, since $k_1 + 1/2$ and $k_2 + 1/2$

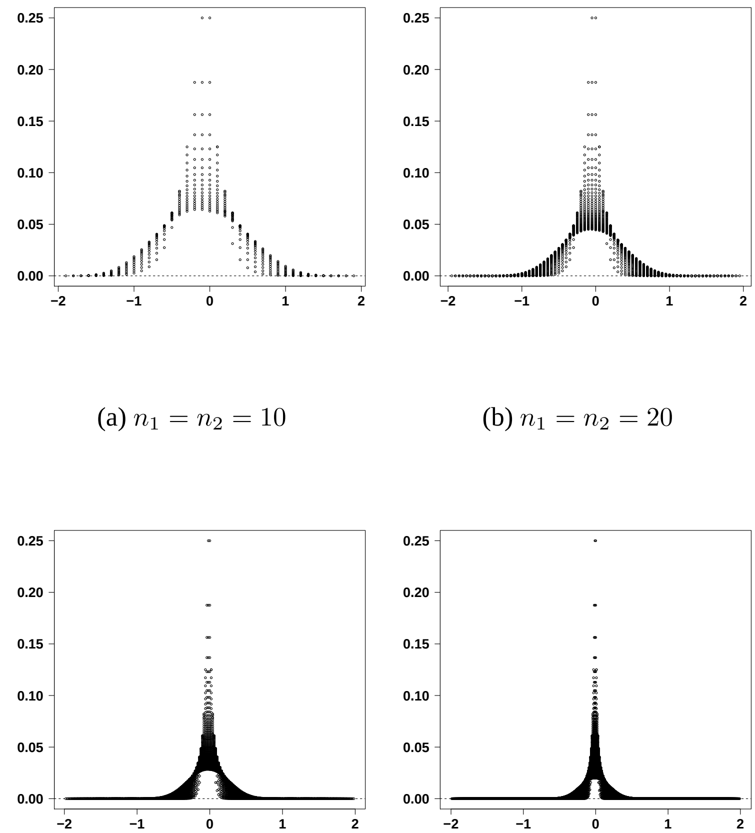
are not natural numbers, $\theta = I_{n_1c/(n_1c+n_2)}(k_1 + 1/2, k_2 + 1/2)$ does not equal $1 - p$.

3.5 Plot of θ and p-value

In this section, we plot and compare the Bayesian index θ and $1 - p$ of the conditional test.

3.5.1 Plot of $\theta = \mathbf{Pr}(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ and the p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$

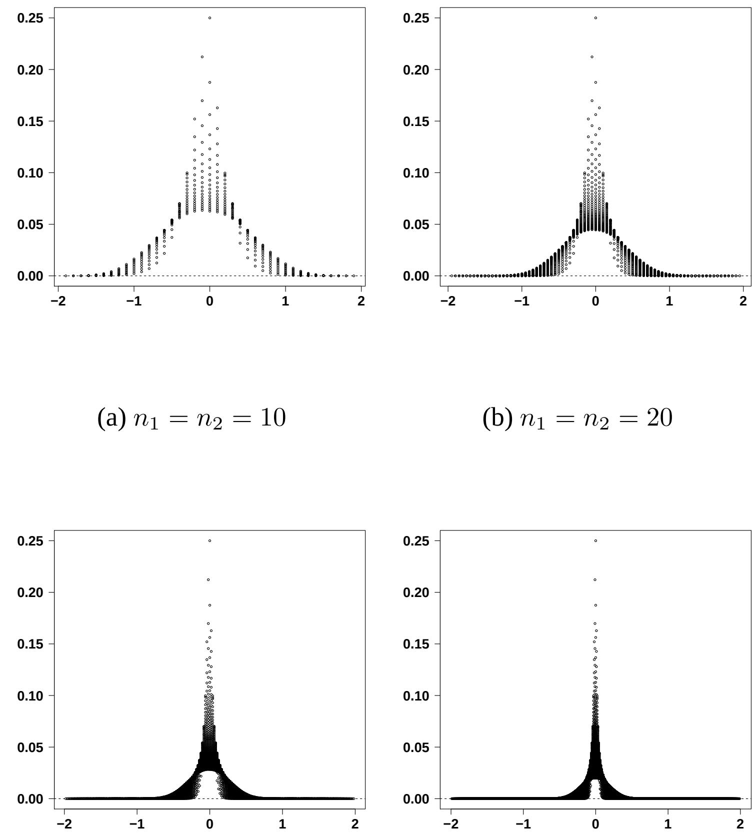
First, we calculate $\theta = \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ and the one-sided p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$ for all pairs of (X_1, X_2) satisfying $1 \leq X_1 \leq 2n_1$ and $1 \leq X_2 \leq 2n_2$, for $n_1 = n_2 = 10, 20, 50,$ and 100, respectively. Here, we consider four types of prior distribution: (i) non-informative prior for both λ_1 and λ_2, (ii) Jeffreys prior for both λ_1 and λ_2, (iii) $Ga(20, 20)$ for both λ_1 and λ_2, (iv) $Ga(20, 20)$ for λ_1 and $Ga(30, 20)$ for λ_2. (iii) is the prior supporting $H_0 : \lambda_1 \geq \lambda_2$, and (iv) is the prior supporting $H_1 : \lambda_1 < \lambda_2$. Figure 3.1 to 3.4 are the results with the prior (i) to (iv), respectively, where the horizontal axises show the differences between sample rates $\widehat{\lambda}_1 - \widehat{\lambda}_2$ where $\widehat{\lambda}_i = X_i/n_i$, and the vertical axises show the differences $\theta - (1 - p)$. Figure 3.1 and 3.2 show that the result with the non-informative prior and Jeffreys prior are similar. Figure 3.3 shows that θ may be less than $1 - p$ when $\widehat{\lambda}_1 - \widehat{\lambda}_2$ is negative and near 0. This indicates that the prior supporting H_0 may decrease $\theta = \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$, which is the posterior probability of H_1 being true. Figure 3.4 shows that θ with the prior supporting H_1 may tend to be much more than $1 - p$ when $\widehat{\lambda}_1 - \widehat{\lambda}_2$ is positive and small. Figure 3.1 to 3.4 show that θ and $1 - p$ are similar when $|\widehat{\lambda}_1 - \widehat{\lambda}_2|$ are large. Furthermore, as larger n_1 and n_2 are, the more similar θ and $1 - p$ are for moderate $|\widehat{\lambda}_1 - \widehat{\lambda}_2|$.



(a) $n_1 = n_2 = 10$ (b) $n_1 = n_2 = 20$

(c) $n_1 = n_2 = 50$ (d) $n_1 = n_2 = 100$

Figure 3.1: The comparison of $\theta - (1-p)$ and $\widehat{\lambda}_1 - \widehat{\lambda}_2$ for $\theta = \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ with the non-informative prior and the p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$ (vertical axis: $\theta - (1-p)$. horizontal axis: $\widehat{\lambda}_1 - \widehat{\lambda}_2$).



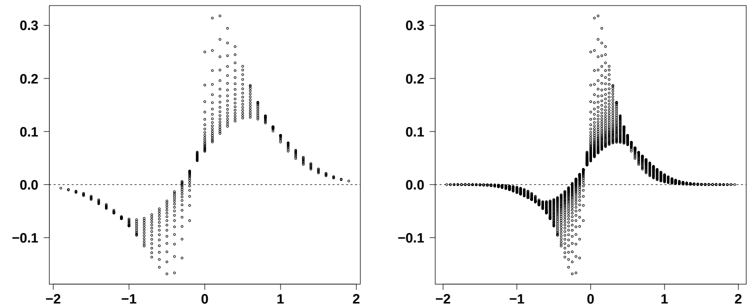
(a) $n_1 = n_2 = 10$

(b) $n_1 = n_2 = 20$

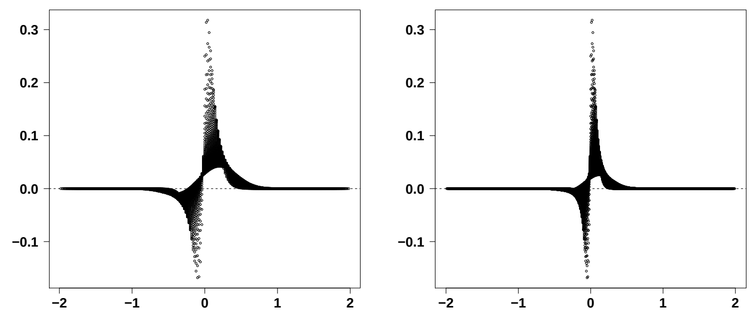
(c) $n_1 = n_2 = 50$

(d) $n_1 = n_2 = 100$

Figure 3.2: The comparison of $\theta - (1 - p)$ and $\widehat{\lambda}_1 - \widehat{\lambda}_2$ for $\theta = \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ with Jeffreys prior and the p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$ (vertical axis: $\theta - (1 - p)$. horizontal axis: $\widehat{\lambda}_1 - \widehat{\lambda}_2$).

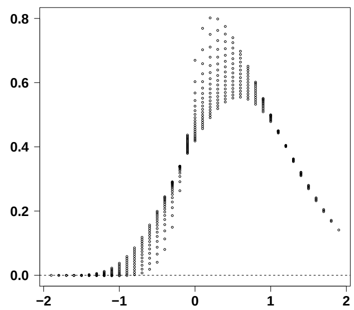


(a) $n_1 = n_2 = 10$ (b) $n_1 = n_2 = 20$

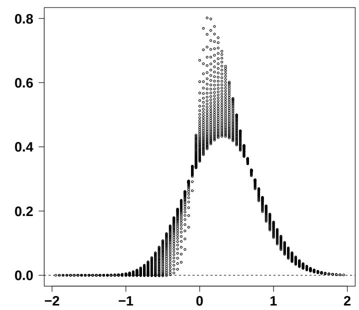


(c) $n_1 = n_2 = 50$ (d) $n_1 = n_2 = 100$

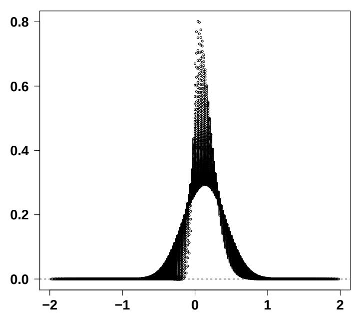
Figure 3.3: The comparison of $\theta - (1 - p)$ and $\widehat{\lambda}_1 - \widehat{\lambda}_2$ for $\theta = \text{Pr}(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ with $Ga(20, 20)$ as the prior for both λ_1 and λ_2 and the p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$ (vertical axis: $\theta - (1 - p)$. horizontal axis: $\widehat{\lambda}_1 - \widehat{\lambda}_2$).



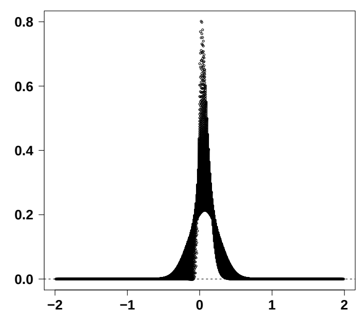
(a) $n_1 = n_2 = 10$



(b) $n_1 = n_2 = 20$



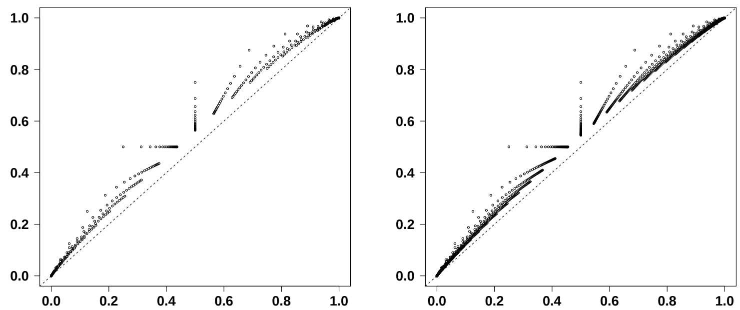
(c) $n_1 = n_2 = 50$



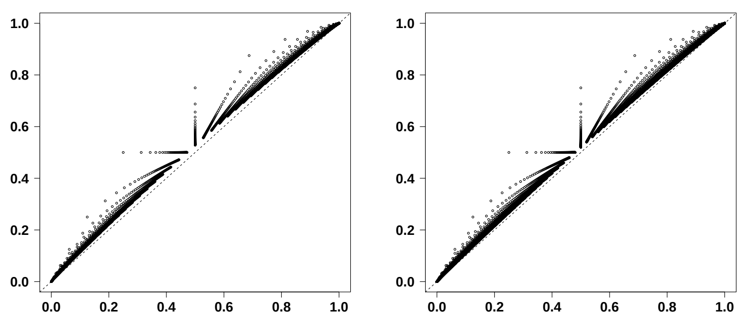
(d) $n_1 = n_2 = 100$

Figure 3.4: The comparison of $\theta - (1-p)$ and $\widehat{\lambda}_1 - \widehat{\lambda}_2$ for $\theta = \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ with $Ga(20, 20)$ as the prior for λ_1 and $Ga(30, 20)$ as the prior for λ_2 and the p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$ (vertical axis: $\theta - (1-p)$. horizontal axis: $\widehat{\lambda}_1 - \widehat{\lambda}_2$).

Figure 3.5 to 3.8 are the results with the prior (i) to (iv), respectively, where the horizontal axises show $1-p$, and the vertical axises show θ. Figure 3.5 and 3.6 show that the result with non-informative prior and Jeffreys prior are similar. Figure 3.7 shows that θ with the prior supporting H_0 can be less than $1-p$ when both of θ and $1-p$ are near 1. Figure 3.8 shows that θ with the prior supporting H_1 tends to be much more than $1-p$.

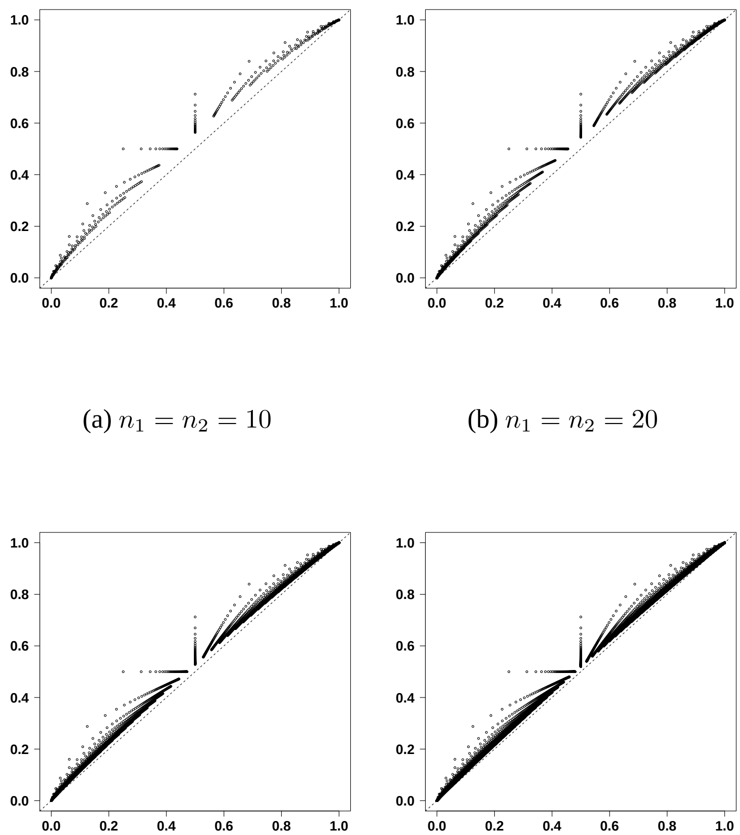


(a) $n_1 = n_2 = 10$ (b) $n_1 = n_2 = 20$



(c) $n_1 = n_2 = 50$ (d) $n_1 = n_2 = 100$

Figure 3.5: The comparison of θ given $X_1 = k_1, X_2 = k_2$ and $1 - p$ given $X_1 = k_1, X_2 = k_2$ for $\theta = \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ with the non-informative prior and the p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$. (vertical axis: θ. horizontal axis: $1 - p$).



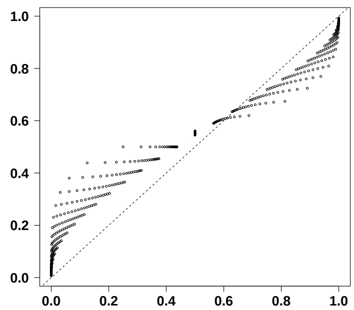
(a) $n_1 = n_2 = 10$

(b) $n_1 = n_2 = 20$

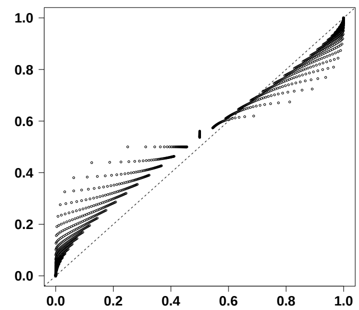
(c) $n_1 = n_2 = 50$

(d) $n_1 = n_2 = 100$

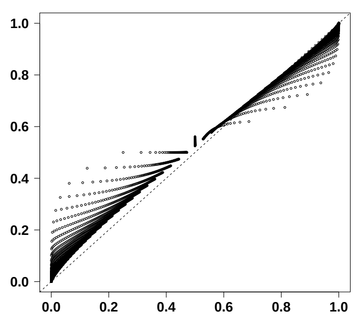
Figure 3.6: The comparison of θ given $X_1 = k_1, X_2 = k_2$ and $1 - p$ given $X_1 = k_1, X_2 = k_2$ for $\theta = \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ with Jeffreys prior and the p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs $H_1 : \lambda_1 < \lambda_2$ (vertical axis: θ. horizontal axis: $1 - p$).



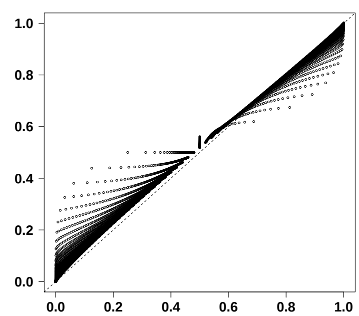
(a) $n_1 = n_2 = 10$



(b) $n_1 = n_2 = 20$

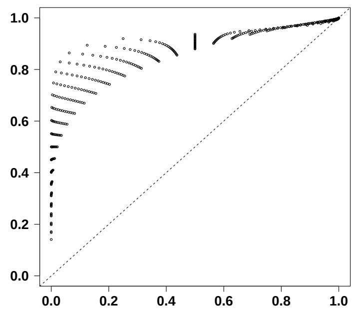


(c) $n_1 = n_2 = 50$

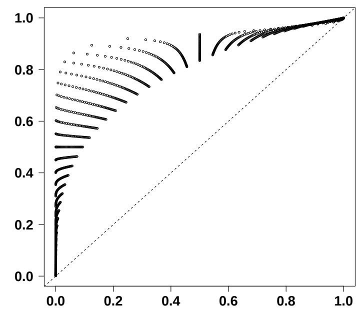


(d) $n_1 = n_2 = 100$

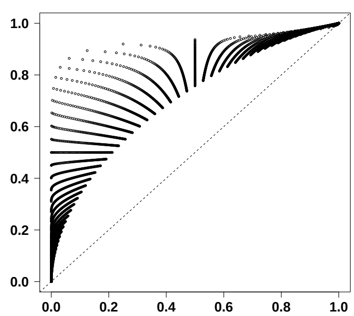
Figure 3.7: The comparison of θ given $X_1 = k_1, X_2 = k_2$ and $1 - p$ given $X_1 = k_1, X_2 = k_2$ for $\theta = \Pr(\lambda_1 < \lambda_2 \mid X_1, X_2)$ with $Ga(20, 20)$ as the prior for both λ_1 and λ_2, and the p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$ (vertical axis: θ. horizontal axis: $1 - p$).



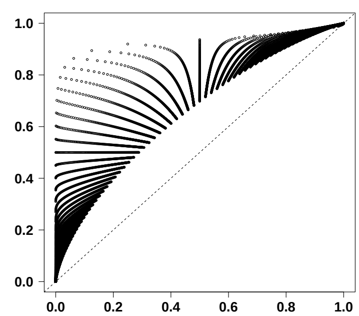
(a) $n_1 = n_2 = 10$



(b) $n_1 = n_2 = 20$



(c) $n_1 = n_2 = 50$

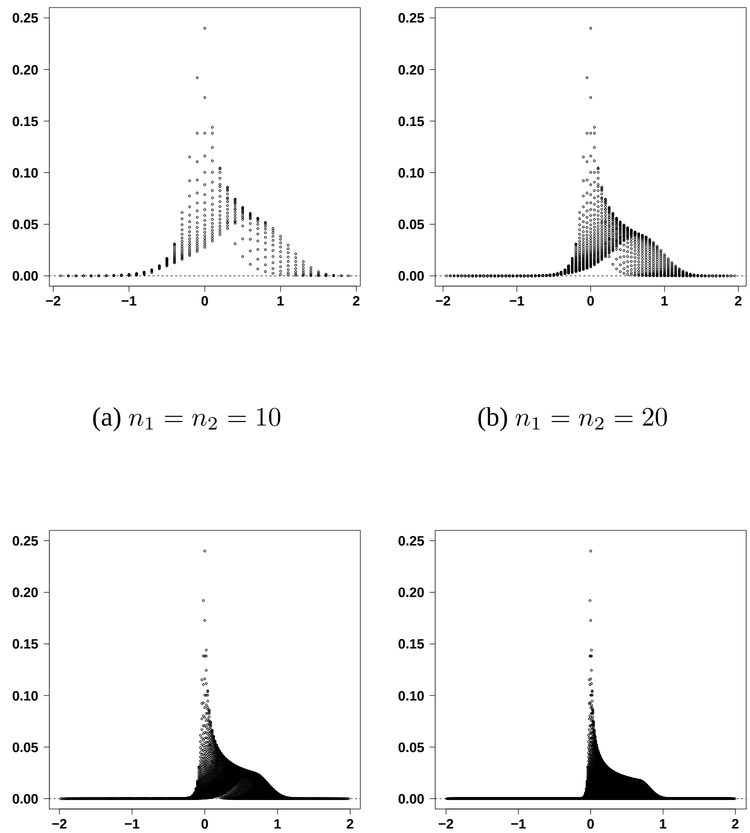


(d) $n_1 = n_2 = 100$

Figure 3.8: The comparison of θ given $X_1 = k_1, X_2 = k_2$ and $1 - p$ given $X_1 = k_1, X_2 = k_2$ for $\theta = \Pr(\lambda_1 < \lambda_2 \,|\, X_1, X_2)$ with $Ga(20, 20)$ as the prior for λ_1 and $Ga(30, 20)$ as the prior for λ_2, and the p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$ (vertical axis: θ. horizontal axis: $1 - p$).

3.5.2 Plot of $\theta = \mathbf{Pr}(\lambda_1/\lambda_2 < c \,|\, X_1, X_2)$ and the p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq c$ vs. $H_1 : \lambda_1/\lambda_2 < c$

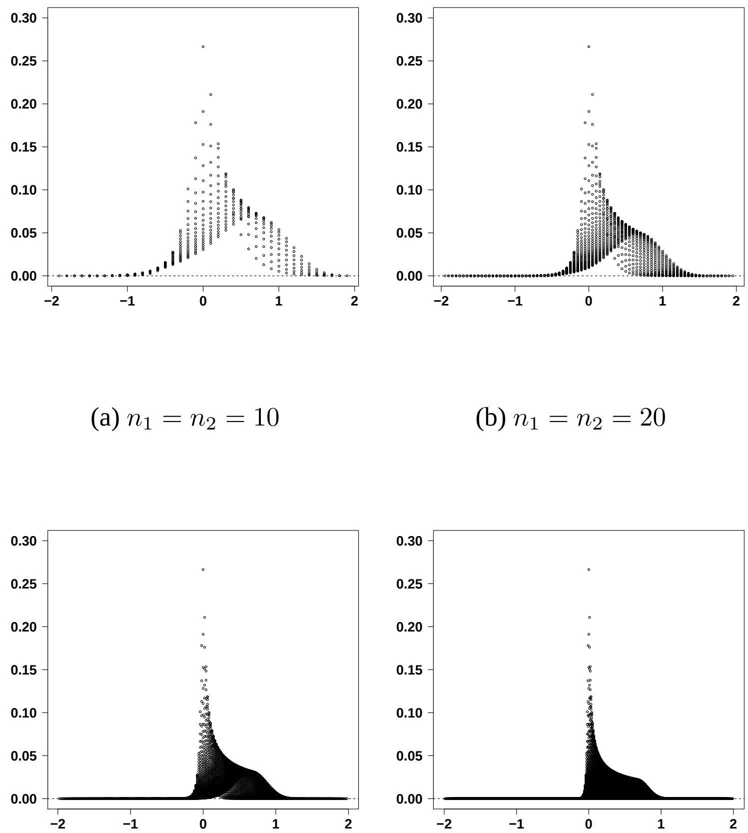
Next, we calculate $\theta = \Pr(\lambda_1/\lambda_2 < c \,|\, X_1, X_2)$ and the one-sided p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq c$ vs. $H_1 : \lambda_1/\lambda_2 < c$ for all pairs of (X_1, X_2) satisfying $1 \leq X_1 \leq 2n_1$ and $1 \leq X_2 \leq 2n_2$, for $n_1 = n_2 = 10, 20, 50,$ and 100, respectively. Then, we take $c = 1.5$ and consider four types of the prior distribution: (i) non-informative prior for both λ_1 and λ_2, (ii) Jeffreys prior for both λ_1 and λ_2, (iii) $Ga(30, 20)$ for λ_1 and $Ga(20, 20)$ for λ_2, (iv) $Ga(20, 20)$ for both λ_1 and λ_2. Here, (iii) is the prior supporting $H_0 : \lambda_1/\lambda_2 \geq c$, and (iv) is the prior supporting $H_1 : \lambda_1/\lambda_2 < c$. Figure 3.9 to 3.12 are the results with the prior (i) to (iv), respectively, where the horizontal axises show the differences between sample rates $\widehat{\lambda}_1 - \widehat{\lambda}_2$ where $\widehat{\lambda}_i = X_i/n_i$, and the vertical axises show the differences $\theta - (1 - p)$. Figure 3.9 to 3.12 are similar to Figure 3.1 to 3.4, respectively, but they show more asymmetry.



(a) $n_1 = n_2 = 10$ (b) $n_1 = n_2 = 20$

(c) $n_1 = n_2 = 50$ (d) $n_1 = n_2 = 100$

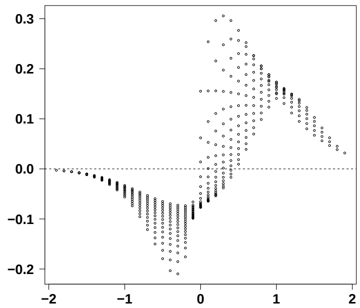
Figure 3.9: The comparison of $\theta - (1 - p)$ and $\widehat{\lambda}_1 - \widehat{\lambda}_2$ for $\theta = \Pr(\lambda_1/\lambda_2 < 1.5 \,|\, X_1, X_2)$ with the non-informative prior and the p-value of the conditional test with H_0 : $\lambda_1/\lambda_2 \geq 1.5$ vs. $H_1 : \lambda_1/\lambda_2 < 1.5$ (vertical axis: $\theta-(1-p)$. horizontal axis: $\widehat{\lambda}_1 - \widehat{\lambda}_2$).



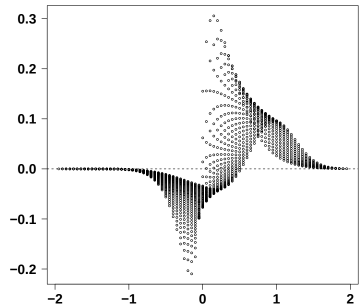
(a) $n_1 = n_2 = 10$ (b) $n_1 = n_2 = 20$

(c) $n_1 = n_2 = 50$ (d) $n_1 = n_2 = 100$

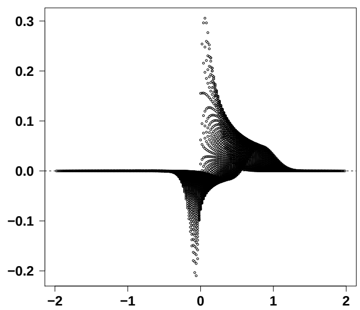
Figure 3.10: The comparison of $\theta - (1 - p)$ and $\widehat{\lambda}_1 - \widehat{\lambda}_2$ for $\theta = \Pr(\lambda_1/\lambda_2 < 1.5 \,|\, X_1, X_2)$ with Jeffreys prior and the p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq 1.5$ vs $H_1 : \lambda_1/\lambda_2 < 1.5$ (vertical axis: $\theta - (1 - p)$. horizontal axis: $\widehat{\lambda}_1 - \widehat{\lambda}_2$).



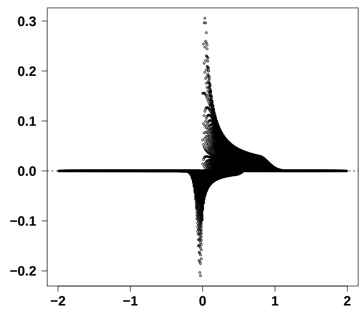
(a) $n_1 = n_2 = 10$



(b) $n_1 = n_2 = 20$

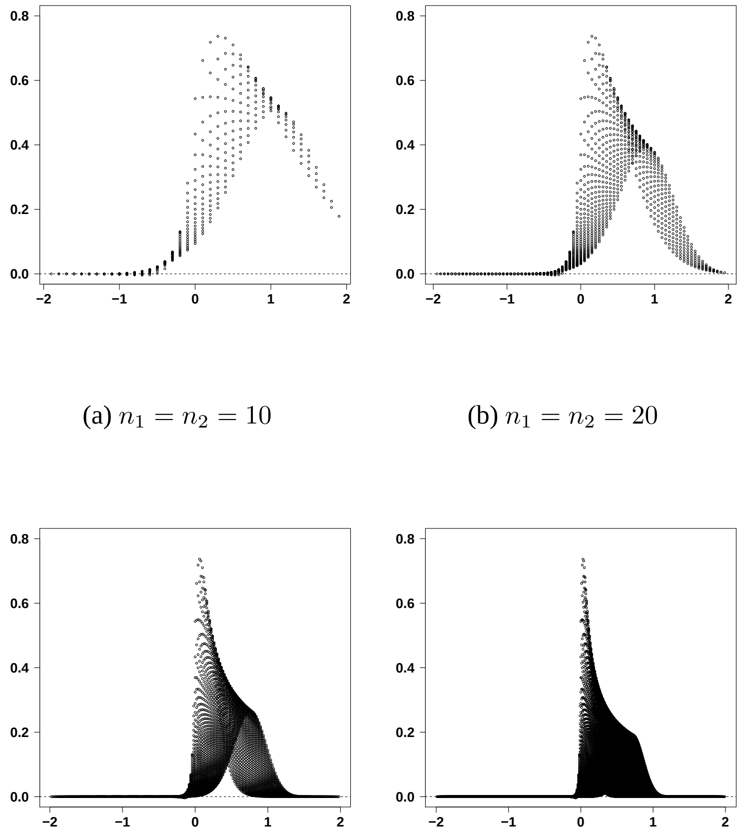


(c) $n_1 = n_2 = 50$



(d) $n_1 = n_2 = 100$

Figure 3.11: The comparison of $\theta - (1 - p)$ and $\widehat{\lambda}_1 - \widehat{\lambda}_2$ for $\theta = \Pr(\lambda_1/\lambda_2 < 1.5 \,|\, X_1, X_2)$ with $Ga(30, 20)$ as the prior for λ_1 and $Ga(20, 20)$ as the prior for λ_2 and the p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq 1.5$ vs $H_1 : \lambda_1/\lambda_2 < 1.5$ (vertical axis: $\theta - (1 - p)$. horizontal axis: $\widehat{\lambda}_1 - \widehat{\lambda}_2$).

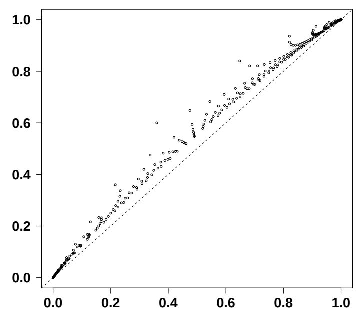


(a) $n_1 = n_2 = 10$ (b) $n_1 = n_2 = 20$

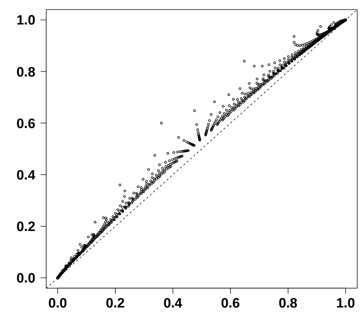
(c) $n_1 = n_2 = 50$ (d) $n_1 = n_2 = 100$

Figure 3.12: The comparison of $\theta - (1-p)$ and $\widehat{\lambda}_1 - \widehat{\lambda}_2$ for $\theta = \Pr(\lambda_1/\lambda_2 < 1.5 \,|\, X_1, X_2)$ with $Ga(20, 20)$ as the prior for both λ_1 and λ_2 and the p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq 1.5$ vs. $H_1 : \lambda_1/\lambda_2 < 1.5$ (vertical axis: $\theta - (1-p)$. horizontal axis: $\widehat{\lambda}_1 - \widehat{\lambda}_2$).

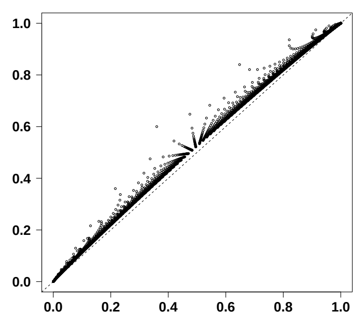
Figure 3.13 to 3.16 show the results with the prior (i) to (iv), respectively, where the horizontal axises show $1 - p$, and the vertical axises show θ. The results are similar to Figure 3.5 to 3.8, respectively.



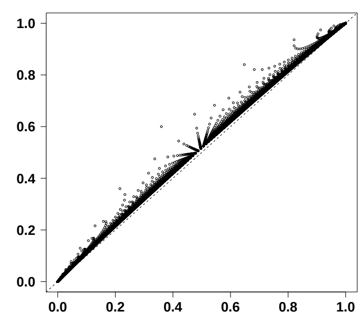
(a) $n_1 = n_2 = 10$



(b) $n_1 = n_2 = 20$

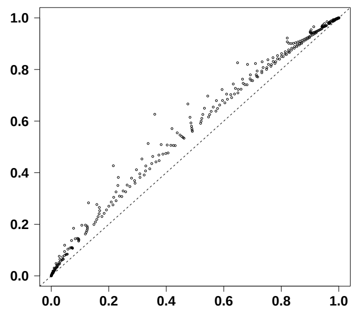


(c) $n_1 = n_2 = 50$

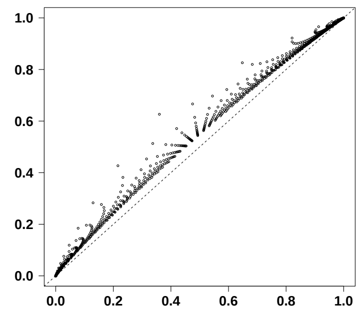


(d) $n_1 = n_2 = 100$

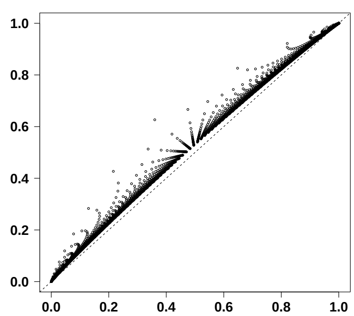
Figure 3.13: The comparison of θ given $X_1 = k_1, X_2 = k_2$ and $1 - p$ given $X_1 = k_1, X_2 = k_2$ for $\theta = \Pr(\lambda_1/\lambda_2 < 1.5 \mid X_1, X_2)$ with the non-informative prior and the p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq 1.5$ vs. $H_1 : \lambda_1/\lambda_2 < 1.5$ (vertical axis: θ. horizontal axis: $1 - p$).



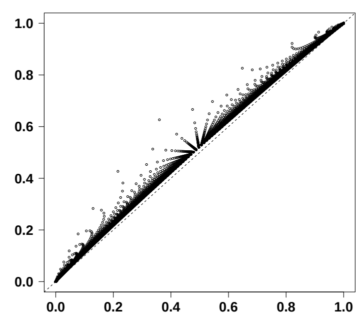
(a) $n_1 = n_2 = 10$



(b) $n_1 = n_2 = 20$

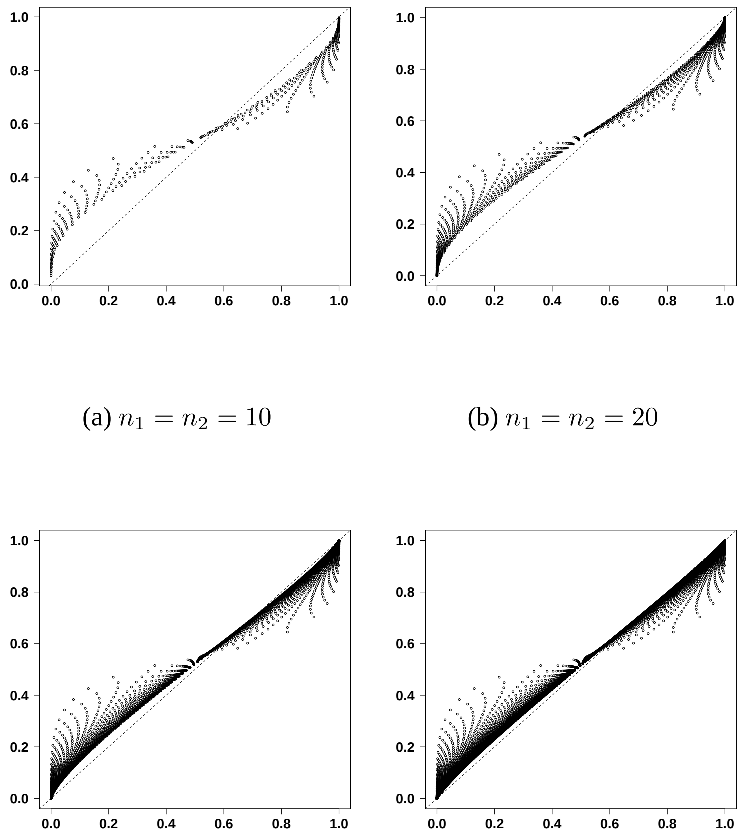


(c) $n_1 = n_2 = 50$



(d) $n_1 = n_2 = 100$

Figure 3.14: The comparison of θ given $X_1 = k_1, X_2 = k_2$ and $1 - p$ given $X_1 = k_1, X_2 = k_2$ for $\theta = \Pr(\lambda_1/\lambda_2 < 1.5 \mid X_1, X_2)$ with Jeffreys prior and the p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq 1.5$ vs. $H_1 : \lambda_1/\lambda_2 < 1.5$ (vertical axis: θ. horizontal axis: $1 - p$).



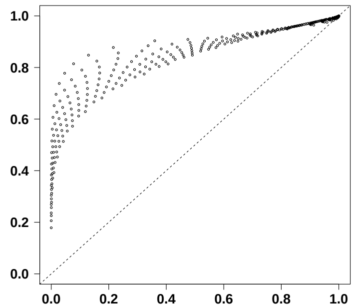
(a) $n_1 = n_2 = 10$

(b) $n_1 = n_2 = 20$

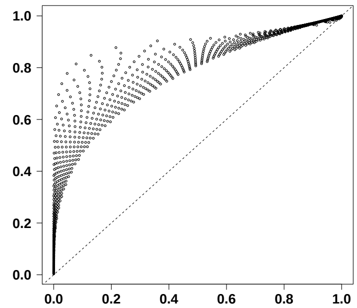
(c) $n_1 = n_2 = 50$

(d) $n_1 = n_2 = 100$

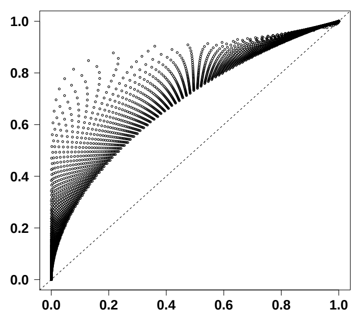
Figure 3.15: The comparison of θ given $X_1 = k_1, X_2 = k_2$ and $1 - p$ given $X_1 = k_1, X_2 = k_2$ for $\theta = \Pr(\lambda_1/\lambda_2 < 1.5 \,|\, X_1, X_2)$ with $Ga(30, 20)$ as the prior for λ_1 and $Ga(20, 20)$ as the prior for λ_2, and the p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq 1.5$ vs. $H_1 : \lambda_1/\lambda_2 < 1.5$ (vertical axis: θ. horizontal axis: $1 - p$).



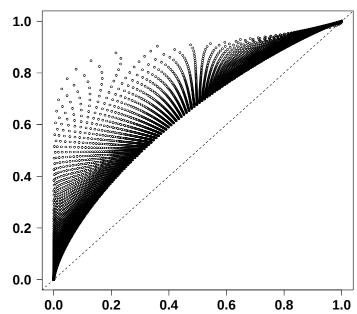
(a) $n_1 = n_2 = 10$



(b) $n_1 = n_2 = 20$



(c) $n_1 = n_2 = 50$



(d) $n_1 = n_2 = 100$

Figure 3.16: The comparison of θ given $X_1 = k_1, X_2 = k_2$ and $1 - p$ given $X_1 = k_1, X_2 = k_2$ for $\theta = \Pr(\lambda_1/\lambda_2 < 1.5 \,|\, X_1, X_2)$ with $Ga(20, 20)$ as the prior for both λ_1 and λ_2, and the p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq 1.5$ vs. $H_1 : \lambda_1/\lambda_2 < 1.5$ (vertical axis: θ. horizontal axis: $1 - p$).

3.6 Application

In this section, we apply the Bayesian index to real epidemiology and clinical trial data and compare it to the one-sided p-value of the conditional test.

Example 4 (Breast cancer study)**.**

Table 3.1 shows the result of a breast cancer study reported in Rothman et al. (2008) . The rates of breast cancer between two groups of women are compared. One group is composed of the women with tuberculosis who are repeatedly exposed to multiple x-ray fluoroscopies and the other group is composed of unexposed women with tuberculosis.

Table 3.1: Breast cancer study data

	cases of breast cancer (X_i)	person-years at risk (n_i)
Received x-ray fluoroscopy $(i = 1)$	41	28,010
Control $(i = 2)$	15	19,017

Let X_1, X_2 be the independent Poisson processes indicating the numbers of cases of breast cancer, and n_1, n_2 be person-years at risk. Here, we suppose $X_i \overset{ind}{\sim} Po(n_i\lambda_i)$ for $i = 1, 2$. From Table 3.1, $X_1 = 41, n_1 = 28,010$ and $X_2 = 15, n_2 = 19,017$. First, we consider the conditional test with the null hypothesis $H_0 : \lambda_1 \leq \lambda_2$ versus the alternative $H_1 : \lambda_1 > \lambda_2$ and the Bayesian index $\theta = \Pr(\lambda_1 > \lambda_2 \,|X_1, X_2, n_1, n_2)$ with the non-informative and Jeffreys priors. Table 3.2 shows the results. Here, $1 - p = 0.976$. Hence, $\theta - (1 - p) = 0.009$ and 0.007 with the non-informative and Jeffreys priors, respectively.

Table 3.2: p-value with $H_0 : \lambda_1 \leq \lambda_2$ vs. $H_1 : \lambda_1 > \lambda_2$ and Bayesian index $\Pr(\lambda_1 > \lambda_2 \,|\, X_1, X_2)$ for the breast cancer study data

p-value with	Bayesian index $\Pr(\lambda_1 > \lambda_2 \,\|\, X_1, X_2)$	
$H_0 : \lambda_1 \leq \lambda_2$ vs. $H_1 : \lambda_1 > \lambda_2$	non-informative prior	Jeffreys prior
0.024	0.985	0.983

Next, as in Gu et al. (2008), we consider the test with the null hypothesis $H_0 : \lambda_1/\lambda_2 \leq 1.5$ versus the alternative $H_1 : \lambda_1/\lambda_2 > 1.5$ and the Bayesian index $\Pr(\lambda_1/\lambda_2 > 1.5 \,|\, X_1, X_2)$ with the non-informative and Jeffreys priors. Table 3.3 shows the results. In this case, $1 - p = 0.709$.

Hence, $\theta - (1-p) = 0.067$ and 0.048 with the non-informative and Jeffreys priors, respectively.

Table 3.3: p-value with $H_0 : \lambda_1/\lambda_2 \leq 1.5$ vs. $H_1 : \lambda_1/\lambda_2 > 1.5$ and Bayesian index $\Pr(\lambda_1/\lambda_2 > 1.5 \mid X_1, X_2)$ for the breast cancer study data

p-value with	Bayesian index $\Pr(\lambda_1/\lambda_2 > 1.5 \mid X_1, X_2)$	
$H_0 : \lambda_1/\lambda_2 \leq 1.5$ vs. $H_1 : \lambda_1/\lambda_2 > 1.5$	non-informative prior	Jeffreys prior
0.291	0.776	0.757

Example 5 (Hypertension trials)**.**

Table 3.4 shows the result of the two selected hypertension clinical trials in Table II in Arends et al. (2000) . We assume that the trial 1 is of interest and we utilize the trial 2 data to specify the conditional power priors described in section 3.2.2 Let X_1, X_2 be the independent Poisson process indicating the number of deaths, and n_1, n_2 be the number of the person-year in trial 1, and let x_1^0, x_2^0 be the independent Poisson process indicating the number of deaths, and m_1, m_2 be the number of the person-year in trial 2. Here we suppose $X_i \sim Po(n_i\lambda_i)$, $x_i^0 \sim Po(m_i\lambda_i)$, and X_1, X_2, x_1^0, x_2^0 are independent. From Table 3.4, $X_1 = 54, n_1 = 5,635,\ X_2 = 70, n_2 = 5,600,\ x_1^0 = 47, m_1 = 5,135,\ x_2^0 = 63, m_2 = 4,960$.

Table 3.4: Hypertension trials data

	Treatment group		Control group	
	death	number of person-year	death	number of person-year
Trial 1	54	5,635	70	5,600
Trial 2	47	5,135	63	4,960

We consider the test with the null hypothesis $H_0 : \lambda_1 \geq \lambda_2$ versus the alternative $H_1 : \lambda_1 < \lambda_2$ and the Bayesian index $\theta = \Pr(\lambda_1 < \lambda_2 \mid X_1, X_2)$ with the non-informative prior and the conditional power priors. For the conditional power prior, we assume $a_1 = a_2 (=: a)$ and take $a = 0.1, 0.5$, and 1.0. Table 3.5 shows the result. Here $1 - p = 0.917$ and θ with the non-informative prior is 0.930. Additionally, θ with the conditional power priors are greater than that with the non-informative prior. Moreover, when a increases, θ also increases.

Table 3.5: p-value and Bayesian index for the hypertension trials data ($a = a_1 = a_2$)

p-value with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$	Bayesian index $\Pr(\lambda_1 < \lambda_2 \mid X_1, X_2)$			
	non-informative prior	conditional power prior		
		$a = 0.1$	$a = 0.5$	$a = 1.0$
0.083	0.930	0.942	0.971	0.988

Here, we treat that $\theta > 0.975$ and $p < 0.025$ show the superiority of treatment from the Bayesian and frequentist perspective, respectively. Both of them equal to each other when $p = 1 - \theta$. Here, the result from θ with the non-informative prior and the result of the conditional test are similar because θ is similar to $1 - p$. On the other hand, with the conditional power prior with $\alpha_1 = \alpha_2 = 1.0$, the result from θ shows the superiority of treatment although the result from the conditional test does not. In this case, by borrowing strength from the historical data, the Bayesian index leads the conclusion that differs from the conditional test. Table 3.6 shows the result of two clinical trials of teriflunomide in Multiple Sclerosis (TEMSO and TOWER trials) in HAS (2014). $i = 1$ indicates the teriflunomide 7mg group, and $i = 2$ indicates the placebo group. We compare the relapse rate λ_1 and λ_2 for each group. Similar to example 5, we assume that TOWER trial is of interest and we utilize the TEMSO trial data to specify the conditional power priors. Let X_1, X_2 be the independent Poisson process indicating the total number of relapses and n_1, n_2 be the total number of patient-years for each group of the TOWER trial, and let x_1^0, x_2^0 be the independent Poisson process indicating the total number of relapses and m_1, m_2 be the total number of patient-years for each group of the TEMSO trial. From Table 3.6, $X_1 = 235, n_1 = 614, X_2 = 296, n_2 = 608, x_1^0 = 233, m_1 = 634, x_2^0 = 335, m_2 = 628$. For simplicity, we round n_1, n_2, m_1, m_2 to the nearest integer.

Table 3.6: MS trials data

	Teriflunomide 7mg group		Placebo group	
	relapse	number of person-year	relapse	number of person-year
TOWER	235	614	296	608
TEMSO	233	634	335	628

We consider the conditional test with the null hypothesis $H_0 : \lambda_1/\lambda_2 \geq c$ versus the alternative $H_1 : \lambda_1/\lambda_2 < c$ and the Bayesian index $\theta = \Pr(\lambda_1/\lambda_2 < c \mid X_1, X_2)$ with the non-informative prior and the conditional power priors for $c = 1.0, 0.9, 0.8$. For the conditional power prior, we assume $a_1 = a_2 (=: a)$ and take $a = 0.1, 0.5, 1.0$. Table 3.7 shows the result.

Table 3.7: p-value and Bayesian index for the MS trials data ($a = a_1 = a_2$)

	p-value with	Bayesian index $\Pr(\lambda_1/\lambda_2 < c \mid X_1, X_2)$			
	$H_0 : \lambda_1/\lambda_2 \geq c$ vs. $H_1 : \lambda_1/\lambda_2 < c$	non-informative prior	conditional power prior		
c			$a = 0.1$	$a = 0.5$	$a = 1.0$
1.0	0.003	0.997	0.999	1.000	1.000
0.9	0.066	0.940	0.963	0.995	1.000
0.8	0.439	0.580	0.643	0.815	0.920

Here, same as example 5, we treat that $\theta > 0.975$ and $p < 0.025$ show the superiority of treatment from the Bayesian and frequentist perspective, respectively. When $c = 1.0$, $\lambda_1/\lambda_2 < c$ is shown for both methods. When $c = 0.9$, $p \geq 0.025$, but $\theta > 0.975$ when $a = 0.5, 1.0$. In this case, the Bayesian index shows $\lambda_1/\lambda_2 < c$ by borrowing strength from the historical data, although the conditional test does not. When $c = 0.8$, $\lambda_1/\lambda_2 < c$ cannot be shown for both methods.

3.7 Conclusion

In this chapter, we provided the cumulative distribution function expressions for the Bayesian index $\theta = \Pr(\lambda_1 < \lambda_2 \mid X_1, X_2)$ for the Poisson parameters, which can be more easily calculated than the hypergeometric series expression in Kawasaki and Miyaoka (2012a) . Next, we showed the relationship between the Bayesian index with the non-informative prior and the one-sided p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ versus $H_1 : \lambda_1 < \lambda_2$. This relationship can be considered as the Poisson distribution counterpart of the relationship between the Bayesian index for binomial proportions and the one-sided p-value of Fisher's exact test in Kawasaki et al. (2014). Additionally, we generalized the Bayesian index to $\theta = \Pr(\lambda_1/\lambda_2 < c \mid X_1, X_2)$, expressed it using the cumulative distribution functions and hypergeometric series, and investigated the relationship between θ and the one-sided p-value of the conditional test with $H_0 : \lambda_1/\lambda_2 \geq c$ versus $H_1 : \lambda_1/\lambda_2 < c$. By the analysis of hypertension trials data, we showed that the Bayesian index $\theta = \Pr(\lambda_1 < \lambda_2 \mid X_1, X_2)$ with the non-informative prior is similar to $1 - p$ of the conditional test, and the Bayesian index

θ with the conditional power prior with $a = 1.0$ is greater than 0.975 although $p \geq 0.025$. This indicates that when we treat $p < 0.025$ and $\theta > 0.975$ as superiority of treatment from the frequentist and Bayesian perspective, respectively, the Bayesian index θ with the non-informative prior showed the similar result as the conditional test, and the Bayesian index θ with informative prior can potentially improve the efficiency of inference. By the analysis of MS trials, we showed the similar result as the analysis of the hypertension data between $\theta = \Pr(\lambda_1/\lambda_2 < 0.9 \mid X_1, X_2)$ and the conditional test with $H_0 : \lambda_1/\lambda_2 \geq 0.9$ versus $H_1 : \lambda_1/\lambda_2 < 0.9$. Further studies are needed for choosing the suitable historical data and suitable values of α_1, α_2 for conditional power prior.

Chapter 4

Exact Bayesian non-inferiority test for Poisson rate parameters and switching to superiority

4.1 Introduction

Non-inferiority trials have gained increasing attention for drug and medical device development, and most statistical analysis methods are based on frequentist approaches. Currently, however, Bayesian approaches have gained attention based on two FDA guidances. The first is the "Non-Inferiority Clinical Trials to Establish Effectiveness" (US Food and Drug Administration and others (2016)), which states that "Bayesian methods that incorporate historical information from past active control studies through the use of prior distributions of model parameters provide an alternative approach to evaluating non-inferiority in the NI trial itself". The other one is the "Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials"(US Food and Drug Administration and others (2010)), which states that "An adaptive Bayesian clinical trial can involve . . . Switching the hypothesis of non-inferiority to superiority or vice-versa" and "For Bayesian hypothesis testing, you may use the posterior distribution to calculate the probability that a particular hypothesis is true, given the observed data". As this type of Bayesian hypothesis testing, Gamalo et al. (2016), Gamalo-Siebers et al. (2016), and Ghosh et al. (2016) considered the Bayesian non-inferiority test for the normal mean, while Gamalo et al. (2011), Zaslavsky (2013), Kawasaki and Miyaoka (2013), Kawasaki et al. (2016), and Doi et al. (2017b) considered it for the binomial probability.

In line with this framework, Bayesian hypothesis testing for Poisson rate parameters can be considered. For the superiority test, Kawasaki and Miyaoka (2012a) evaluated the posterior probability $\Pr(\lambda_1 < \lambda_2 \mid X_1, X_2)$ for the Poisson rate parameters λ_1 and λ_2 given the data X_1, X_2. They expressed the posterior probability by using the hypergeometric series, and Doi (2016) showed the relationship between this posterior probability and the one-sided p-value of the conditional test.

Similarly, the Bayesian non-inferiority test can be considered based on the probability $\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2)$ with non-inferiority margin Δ. Kawasaki et al. (2016) evaluated this posterior probability by the normal approximation and Monte Carlo integral. However, the normal approximation is not accurate and the Monte Carlo integral is not deterministic.

Here, we first derive the exact expression for $\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2)$ under mild conditions. The expression is not only exact but also simple and easily computable. Then, we derive an approach that flexibly incorporates prior information. Next, we show the relationship between the Bayesian non-inferiority probability and superiority probability, and the relationship between the posterior probability and the one-sided p-value of the superiority conditional test. Based on this relationship, after we show the non-inferiority, we can naturally consider the superiority test which is consistent with the frequentist test. For this type of study, EMA "Points to consider on switching between superiority and non-inferiority" Committee for Proprietary Medicinal Products (CPMP) and others (2000) states that switching the objective of a trial from non-inferiority to superiority is feasible provided that certain conditions are met.

This chapter is organized as follows. In Section 4.2, we briefly summarize the Bayesian non-inferiority test, Bayesian superiority test, and frequentist superiority conditional test. In Section 4.3, first, we derive the exact formula for the Bayesian non-inferiority probability. Next, we show the relationship between the Bayesian non-inferiority probability and superiority probability, and show the relationship between the Bayesian non-inferiority probability and the one-sided p-value of the superiority conditional test. In Section 4.4, we describe the operating characteristics of the Bayesian non-inferiority test based on Monte Carlo simulations. In Section 4.5, we provide an analysis of the real non-inferiority trial data for the relapsing-remitting multiple sclerosis trials in the switching from the non-inferiority to superiority framework. Finally, we provide concluding remarks in Section 4.6.

4.2 Bayesian and frequentist test for Poisson parameters

In this chapter, we consider two situations. In each case, let $i = 1$ be the test drug group and $i = 2$ be the active control group. For the first situation, let X_{ij} be the outcome of the jth subject in the ith group for $i = 1, 2$ and $j = 1, \dots, n_i$. Suppose X_{ij} independently follows a Poisson distribution $Po(\lambda_i)$ for $i = 1, 2$, and let $X_i = \sum_{j=1}^{n_i} X_{ij}$. For the second case, let X_i be the independent Poisson process with Poisson rate λ_i and let n_i be the total person-years at risk for the ith group for $i = 1, 2$. In each case, $X_i \stackrel{ind}{\sim} Po(n_i\lambda_i)$. In the following, let n_1, n_2 be the fixed integers for simplicity, and we suppose that smaller values of λ_i are preferable.

4.2.1 Bayesian non-inferiority test

First, we consider the Bayesian non-inferiority test. For the moment, we suppose that the prior distribution of λ_i is a gamma distribution $Ga(\alpha_i, \beta_i)$ for $i = 1, 2$. Next, given $X_i = k_i$, the posterior distribution of λ_i is $Ga(a_i, b_i)$ where $a_i = \alpha_i + k_i, b_i = \beta_i + n_i$. In this setting, we evaluate $\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2)$ with a pre-specified non-inferiority margin $\Delta(> 0)$. This can be considered as the Bayesian hypothesis testing stated in the FDA guidance US Food and Drug Administration and others (2010), as it states that "For Bayesian hypothesis testing, you may use the posterior distribution to calculate the probability that a particular hypothesis is true, given the observed data", and for the frequentist null and alternative hypothesis $H_0 : \lambda_1 \geq \lambda_2 + \Delta$ and $H_1 : \lambda_1 < \lambda_2 + \Delta$, the posterior probability above is the probability of H_1 being true given X_1, X_2.

4.2.2 Bayesian and frequentist superiority test

For the frequentist superiority test, we consider the conditional test discussed in Chapter 3. Then, for the relationship between Bayesian and frequentist superiority test, Theorem 3.3 and Corollary 3.4 are important.

4.3 Exact formula and the relationship between superiority and non-inferiority test

4.3.1 Exact formula for posterior probability and relationship between Bayesian non-inferiority and superiority probability

To evaluate the posterior probability $\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2)$, Kawasaki et al. (2016) utilized the normal approximation and the Monte Carlo integral. However, normal approximation is not accurate and Monte Carlo integral is not deterministic. Here, under mild conditions, we derive the exact expression.

Theorem 4.1. If $a_1, a_2 \in \mathbb{N}$, the following relationship holds:

$$\begin{aligned}
&\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) \\
=&1 - \sum_{j=0}^{a_1-1} \frac{(b_1\Delta)^j}{j!} \cdot \exp(-b_1\Delta) \sum_{r=0}^{a_1-1-j} \frac{\Gamma(r+a_2)}{r! \cdot \Gamma(a_2)} \left(\frac{b_2}{b_1+b_2}\right)^{a_2} \left(\frac{b_1}{b_1+b_2}\right)^{r} \\
=&1 - f_{Poi} * F_{NB}(a_1 - 1), \qquad (4.1)
\end{aligned}$$

where

$$f_{Poi}(x) = \begin{cases} 0 & (x = -1, -2, \ldots) \\ \dfrac{(b_1\Delta)^x}{x!} \cdot \exp(-b_1\Delta) & (x = 0, 1, 2, \ldots) \end{cases}$$

is the probability function of the Poisson distribution $Po(b_1\Delta)$,

$$F_{NB}(x) = \begin{cases} 0 & (x = -1, -2, \ldots) \\ \sum_{r=0}^{x} \dfrac{\Gamma(r+a_2)}{r! \cdot \Gamma(a_2)} \left(\dfrac{b_2}{b_1+b_2}\right)^{a_2} \left(\dfrac{b_1}{b_1+b_2}\right)^{r} & (x = 0, 1, 2, \ldots) \end{cases}$$

is the cumulative distribution function of the negative binomial distribution $NB\left(a_2, b_1/(b_1+b_2)\right)$, and $*$ indicates the convolution of sequences, i.e., for two series $f(n)$ and $g(n)$, $f * g(n) = \sum_{m=-\infty}^{\infty} f(n-m) \cdot g(m)$.

Proof. Since $\lambda_i \sim Ga(a_i, b_i)$ and $a_i \in \mathbb{N}$ for $i = 1, 2$,

$$
\begin{aligned}
&\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) \\
=& \int_0^\infty \left(\int_0^{\lambda_2+\Delta} \frac{b_1^{a_1}}{\Gamma(a_1)} \lambda_1^{a_1-1} \exp(-b_1\lambda_1) d\lambda_1 \right) \frac{b_2^{a_2}}{\Gamma(a_2)} \lambda_2^{a_2-1} \exp(-b_2\lambda_2) d\lambda_2 \\
=& \int_0^\infty \left(1 - \int_{\lambda_2+\Delta}^{\infty} \frac{b_1^{a_1}}{\Gamma(a_1)} \lambda_1^{a_1-1} \exp(-b_1\lambda_1) d\lambda_1 \right) \frac{b_2^{a_2}}{\Gamma(a_2)} \lambda_2^{a_2-1} \exp(-b_2\lambda_2) d\lambda_2 \\
=& 1 - \int_0^\infty \left(\int_{\lambda_2+\Delta}^{\infty} \frac{b_1^{a_1}}{\Gamma(a_1)} \lambda_1^{a_1-1} \exp(-b_1\lambda_1) d\lambda_1 \right) \frac{b_2^{a_2}}{\Gamma(a_2)} \lambda_2^{a_2-1} \exp(-b_2\lambda_2) d\lambda_2. \qquad (4.2)
\end{aligned}
$$

First, from 8.352-4 Zwillinger (2014): $\int_x^\infty e^{-t} t^{n-1} dt = (n-1)! \cdot e^{-x} \sum_{m=0}^{n-1} x^m / m!$ for $n \in \mathbb{N}, x \in \mathbb{R}$,

$$
\begin{aligned}
&\int_{\lambda_2+\Delta}^{\infty} \frac{b_1^{a_1}}{\Gamma(a_1)} \lambda_1^{a_1-1} \exp(-b_1\lambda_1) d\lambda_1 \\
=& \sum_{r_1=0}^{a_1-1} \frac{\{b_1(\lambda_2+\Delta)\}^{r_1}}{r_1!} \exp(-b_1(\lambda_2+\Delta)) \\
=& \exp(-b_1\Delta) \sum_{r_1=0}^{a_1-1} \frac{b_1^{r_1}}{r_1!} \sum_{r_2=0}^{r_1} \binom{r_1}{r_2} \lambda_2^{r_2} \Delta^{r_1-r_2} \exp(-b_1\lambda_2). \qquad (4.3)
\end{aligned}
$$

Next, from (4.3),

$$
\begin{aligned}
&\int_0^\infty \left(\int_{\lambda_2+\Delta}^{\infty} \frac{b_1^{a_1}}{\Gamma(a_1)} \lambda_1^{a_1-1} \exp(-b_1\lambda_1) d\lambda_1 \right) \frac{b_2^{a_2}}{\Gamma(a_2)} \lambda_2^{a_2-1} \exp(-b_2\lambda_2) d\lambda_2 \\
=& \exp(-b_1\Delta) \sum_{r_1=0}^{a_1-1} \frac{b_1^{r_1}}{r_1!} \sum_{r_2=0}^{r_1} \binom{r_1}{r_2} \Delta^{r_1-r_2} \cdot \frac{b_2^{a_2}}{\Gamma(a_2)} \int_0^\infty \lambda_2^{r_2+a_2-1} \exp(-(b_1+b_2)\lambda_2) d\lambda_2. \qquad (4.4)
\end{aligned}
$$

Here, let $\lambda := (b_1+b_2)\lambda_2$, then

$$
\begin{aligned}
(4.4) =& \exp(-b_1\Delta) \sum_{r_1=0}^{a_1-1} \frac{b_1^{r_1}}{r_1!} \sum_{r_2=0}^{r_1} \binom{r_1}{r_2} \Delta^{r_1-r_2} \\
&\times \frac{b_2^{a_2}}{\Gamma(a_2)} \int_0^\infty \left(\frac{\lambda}{b_1+b_2} \right)^{r_2+a_2-1} \exp(-\lambda) \cdot \frac{d\lambda}{b_1+b_2} \\
=& \exp(-b_1\Delta) \sum_{r_1=0}^{a_1-1} \frac{b_1^{r_1}}{r_1!} \sum_{r_2=0}^{r_1} \frac{r_1!}{r_2!(r_1-r_2)!} \cdot \Delta^{r_1-r_2} \frac{b_2^{a_2}}{\Gamma(a_2)} \left(\frac{1}{b_1+b_2} \right)^{r_2+a_2} \cdot \Gamma(r_2+a_2)
\end{aligned}
$$

$$= \exp(-b_1\Delta) \sum_{r_1=0}^{a_1-1} \sum_{r_2=0}^{r_1} \frac{(b_1\Delta)^{r_1-r_2}}{(r_1-r_2)!} \frac{\Gamma(r_2+a_2)}{r_2! \cdot \Gamma(a_2)} \left(\frac{b_2}{b_1+b_2}\right)^{a_2} \left(\frac{b_1}{b_1+b_2}\right)^{r_2}. \tag{4.5}$$

Furthermore, let $j = r_1 - r_2$, then

$$\begin{aligned}
(4.5) =& \exp(-b_1\Delta) \sum_{j=0}^{a_1-1} \sum_{r_2=0}^{a_1-1-j} \frac{(b_1\Delta)^j}{j!} \frac{\Gamma(r_2+a_2)}{r_2! \cdot \Gamma(a_2)} \left(\frac{b_2}{b_1+b_2}\right)^{a_2} \left(\frac{b_1}{b_1+b_2}\right)^{r_2} \\
=& \sum_{j=0}^{a_1-1} \frac{(b_1\Delta)^j}{j!} \cdot \exp(-b_1\Delta) \cdot F_{NB}(a_1-1-j) \\
=& f_{Poi} * F_{NB}(a_1-1).
\end{aligned}$$

Finally, from (4.2),

$$\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) = 1 - f_{Poi} * F_{NB}(a_1-1).$$

□

In this expression, we can easily calculate the exact value of the posterior probability. A sample SAS program is provided in 4.6.

Next, consider the relationship between the Bayesian non-inferiority and superiority probability.

Corollary 4.2. For the above setting, if $a_1, a_2 \in \mathbb{N}$, then the following relationship holds

$$\lim_{\Delta \to +0} \Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) = \Pr(\lambda_1 < \lambda_2 \mid X_1, X_2).$$

Proof. Since

$$\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) = 1 - \sum_{j=0}^{a_1-1} \frac{(b_1\Delta)^j}{j!} \cdot \exp(-b_1\Delta) \cdot F_{NB}(a_1-1-j)$$

is the continuous function of Δ,

$$\lim_{\Delta \to +0} \Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) = 1 - F_{NB}(a_1-1) \tag{4.6}$$

$$
\begin{aligned}
&= 1 - \sum_{r=0}^{a_1-1} \frac{\Gamma(r+a_2)}{r! \cdot \Gamma(a_2)} \left(\frac{b_2}{b_1+b_2}\right)^{a_2} \left(\frac{b_1}{b_1+b_2}\right)^{r} \\
&= 1 - I_{b_2/(b_1+b_2)}(a_2, a_1) \quad (\because (3.3), (3.6)) \\
&= I_{b_1/(b_1+b_2)}(a_1, a_2) \quad (\because 26.5.2 \text{ Abramowitz and Stegun (1964)}) \\
&= \Pr(\lambda_1 < \lambda_2 \mid X_1, X_2). \quad (\because (3.3))
\end{aligned}
$$

□

4.3.2 Relationship between Bayesian non-inferiority test and frequentist superiority test

Next, we consider the relationship between the Bayesian non-inferiority probability and the one-sided p-value of the superiority conditional test. In this section, we suppose that historical data exist for both treatment groups as the general situation.

For $i = 1, 2$, let n_{0i} be the sample size or the total person-years of the historical trial, and let $X_{0i} \sim Po(n_{0i}\lambda_i)$ be the historical trial data. Additionally, considering Corollary 3.4, let $f_{01}(\lambda_1) \propto 1, f_{02}(\lambda_2) \propto \lambda_2^{-1}$ be the priors for the historical trial. Then the conditional power priors proposed by Ibrahim and Chen (2000) can be derived as follows

$$
\begin{aligned}
\widetilde{f}(\lambda_1) &\propto L(\lambda_1 \mid X_{01})^{a_{01}} \cdot f_{01}(\lambda_1) \\
&\propto \lambda_1^{(a_{01}X_{01}+1)-1} \exp(-a_{01}n_{01}\lambda_1), \\
\widetilde{f}(\lambda_2) &\propto L(\lambda_1 \mid X_{02})^{a_{02}} \cdot f_{02}(\lambda_2) \\
&\propto \lambda_2^{a_{02}X_{02}-1} \exp(-a_{02}n_{02}\lambda_2),
\end{aligned}
$$

where $0 \le a_{01}, a_{02} \le 1$ are fixed parameters. Then, the conditional power prior of λ_1 is $Ga(a_{01}X_{01}+1, a_{01}n_{01})$ if $a_{01}n_{01} > 0$, and that of λ_2 is $Ga(a_{02}X_{02}, a_{02}n_{02})$ if $a_{02}X_{02} > 0$ and $a_{02}n_{02} > 0$, otherwise $f(\lambda_2) \propto \lambda_2^{-1} \exp(-a_{02}n_{02}\lambda_2)$ or $\propto \lambda_2^{[a_{02}X_02]-1}$. However, as we cannot apply Theorem 4.1 for general gamma priors, we modify these gamma priors. We let

$$
\alpha_1 = [a_{01}X_{01}] + 1, \beta_1 = a_{01}n_{01},
$$

$$\alpha_2 = [a_{02} X_{02}], \beta_2 = a_{02} n_{02}.$$

Further, if $a_{01} n_{01} > 0$ and $[a_{02} X_{02}] > 0$, we suppose the prior of λ_i is $Ga(\alpha_i, \beta_i)$ for $i = 1, 2$ for the present trial.

Next, we show the relationship between the Bayesian posterior non-inferiority probability and the one-sided p-value of the superiority conditional test.

Theorem 4.3.

(i) In the above setting, if $X_2 > 0$, then the following relationship holds

$$\lim_{a_{01}, a_{02}, \Delta \to +0} \Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) = 1 - p,$$

where p is the one-sided p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$.

(ii) Suppose that the prior of λ_i is $Ga(\alpha_i, \beta_i)$ with $\alpha_i, \beta_i \in \mathbb{N}$ for $i = 1, 2$, and let $m_1, m_2 \in \mathbb{N}$. Then between $\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2)$ given $X_1 = k_1 - \alpha_1 + 1, X_2 = k_2 - \alpha_2, n_1 = m_1 - \beta_1, n_2 = m_2 - \beta_2$ and the one-sided p-value of the conditional test with $H_0 : \lambda_1 \geq \lambda_2$ vs. $H_1 : \lambda_1 < \lambda_2$ given $X_1 = k_1, X_2 = k_2, n_1 = m_1, n_2 = m_2$, the following relationship holds

$$\lim_{\Delta \to +0} \Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) = 1 - p.$$

Proof. (i) First, it is obvious that

$$\begin{aligned} &\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) \\ =&1 - \sum_{j=0}^{a_1 - 1} \frac{(b_1 \Delta)^j}{j!} \cdot \exp(-b_1 \Delta) \cdot F_{NB}(a_1 - 1 - j) \end{aligned}$$

is the continuous function of $a_1, b_1, a_2, b_2, \Delta$, and

$$\begin{aligned} a_1 &= \alpha_1 + k_1 = [a_{01} X_{01}] + 1 + k_1, b_1 = \beta_1 + n_1 = a_{01} n_{01} + n_1, \\ a_2 &= \alpha_2 + k_2 = [a_{02} X_{02}] + k_2, b_2 = \beta_2 + n_2 = a_{02} n_{02} + n_2. \end{aligned}$$

Therefore,

$$\lim_{\alpha_{01} \to +0} (a_1, b_1) = (k_1 + 1, n_1),$$
$$\lim_{\alpha_{02} \to +0} (a_2, b_2) = (k_2, n_2).$$

The right handed sides are the parameters of the posteriors of λ_1, λ_2, where the priors are $f(\lambda_1) \propto 1, f(\lambda_2) \propto \lambda_2^{-1}$, respectively.

Therefore,

$$\begin{aligned}
&\lim_{a_{01}, a_{02}, \Delta \to +0} \Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) \\
&= \lim_{a_{01}, a_{02} \to +0} \{1 - F_{NB}(a_1 - 1)\} \quad (\because (4.6)) \\
&= 1 - \lim_{a_{01}, a_{02} \to +0} \sum_{r=0}^{a_1 - 1} \frac{\Gamma(r + a_2)}{r! \cdot \Gamma(a_2)} \left(\frac{b_2}{b_1 + b_2} \right)^{a_2} \left(\frac{b_1}{b_1 + b_2} \right)^{r} \\
&= 1 - \sum_{r=0}^{k_1} \frac{\Gamma(r + k_2)}{r! \cdot \Gamma(k_2)} \left(\frac{n_2}{n_1 + n_2} \right)^{k_2} \left(\frac{n_1}{n_1 + n_2} \right)^{r} \\
&= 1 - \sum_{r=0}^{k_1} \binom{k_1 + k_2}{r} \left(\frac{n_1}{n_1 + n_2} \right)^{r} \left(\frac{n_2}{n_1 + n_2} \right)^{k_1 + k_2 - r} \\
&\quad (\because (3.6)) \\
&= 1 - p. \quad (\because (3.8))
\end{aligned}$$

(ii) The proof directly follows from Theorem 3.3 (ii), Theorem 4.1 and (3.8). □

Using this theorem, we can consider our Bayesian non-inferiority test as the non-inferiority extension of the frequentist superiority conditional test. Combining this relationship with Corollary 4.2, we can consider switching from the non-inferiority to superiority test, which is consistent with the frequentist conditional test. The switching procedure is as follows:

(i) If $\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) > 0.975$ holds, non-inferiority is demonstrated.

(ii) Furthermore, if $\Pr(\lambda_1 < \lambda_2 \mid X_1, X_2) > 0.975$ holds, superiority is also demonstrated.

Step (ii) is consistent with the one-sided superiority conditional test with 2.5% confidence level when the priors are $f(\lambda_1) \propto 1, f(\lambda_2) \propto \lambda_2^{-1}$.

4.4 Simulations

In this section, we evaluate the operating characteristics of our Bayesian non-inferiority test via Monte Carlo simulations. Here, following FDA guidance US Food and Drug Administration and others (2016), which states "Bayesian methods that incorporate historical information from past active control studies through the use of prior distributions of model parameters provide an alternative approach to evaluating non-inferiority in the NI trial itself", we suppose that the historical trial data exist only for the active control group.

We treat λ_1, λ_2 as random variables from the Bayesian viewpoint, and $\bar{\lambda}_1, \bar{\lambda}_2$ as the fixed parameters for generating simulated data. The null and alternative hypotheses are $H_0 : \bar{\lambda}_1 \geq \bar{\lambda}_2 + \Delta$ and $H_1 : \bar{\lambda}_1 < \bar{\lambda}_2 + \Delta$, respectively, with fixed Δ. Next, we define the type I error rate as the probability satisfying $\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) > 0.975$ under $\bar{\lambda}_1 = \bar{\lambda}_2 + \Delta$, which is included in H_0, and the power as the probability satisfying $\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) > 0.975$ under some situations included in H_1.

The simulation procedure is as follows:

1. Specify the number of iteration N, Δ, (n_{02}, n_1, n_2), a_{02}, and $(\bar{\lambda}_1, \bar{\lambda}_2)$. Set COUNT = 0.

2. Generate the historical trial data $X_{02} \sim Po(n_{02}\bar{\lambda}_2)$.

3. For $\alpha_2 = [a_{02}X_{02}]$ and $\beta_2 = a_{02}n_{02}$, let the prior distribution of λ_2 be $Ga(\alpha_2, \beta_2)$ if $\alpha_2, \beta_2 > 0$, else let it be proportional to $\lambda_2^{\alpha_2 - 1}$ or $\lambda_2^{-1} \exp(-\beta_2 \lambda_2)$. Let the prior of λ_1 be proportional to 1.

4. Generate the present trial data $X_1 \sim Po(n_1\bar{\lambda}_1)$ and $X_2 \sim Po(n_2\bar{\lambda}_2)$ independently, and derive the posterior distribution of λ_1 and λ_2, respectively. In the following, suppose $X_2 > 0$.

5. From the posterior distribution of λ_1 and λ_2, calculate the posterior probability $\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2)$.

6. If $\Pr(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2) > 0.975$, then increase the COUNT by 1; otherwise, by 0.

7. Go back to step 2. and repeat the simulation for N times.

8. Calculate type I error rate or power by COUNTS$/N$.

The parameters are as follows:

- $N = 10,000, \Delta = 0.10$.
- $a_{02} = 0, 0.1, 0.25, 0.5, 0.75, 1$.
- Sample sizes or total person-years:
  - Balanced 1 $(n_1 = n_2)$: $(n_{02}, n_1, n_2) = (100, 100, 100)$, $(200, 200, 200)$, $(500, 500, 500)$, $(1000, 1000, 1000)$.
  - Balanced 2 $(n_1 = n_2)$: $(n_{02}, n_1, n_2) = (50, 100, 100), (100, 200, 200), (250, 500, 500), (500, 1000, 1000)$.
  - Unbalanced $(n_1 > n_2)$: $(n_{02}, n_1, n_2) = (100, 100, 50), (200, 200, 100), (500, 500, 250), (1000, 1000, 500)$.
- Poisson rate parameters:
  - (Type I error rate) $\bar{\lambda}_1 = 0.4$ and $\bar{\lambda}_2 = 0.3$.
  - (Power) $\bar{\lambda}_1 = 0.3$ and $\bar{\lambda}_2 = 0.3$.

Table 4.1: Type I error rate of the Bayesian method under $\bar{\lambda}_1 = \bar{\lambda}_2 + \Delta$ in $H_0 : \bar{\lambda}_1 < \bar{\lambda}_2 + \Delta$ $(\bar{\lambda}_1 = 0.4, \bar{\lambda}_2 = 0.3, \Delta = 0.1)$.

				a_{02}					
Scenario	n_{02}	n_1	n_2	0.0	0.1	0.25	0.5	0.75	1.0
Balaneced 1	100	100	100	1.98	1.64	1.59	1.70	1.76	2.00
	200	200	200	2.07	1.81	1.68	1.80	1.84	2.17
	500	500	500	2.26	2.02	1.98	2.07	2.19	2.43
	1000	1000	1000	2.37	2.16	2.07	2.15	2.25	2.50
Balanced 2	50	100	100	1.99	1.69	1.67	1.72	1.79	2.02
	100	200	200	2.04	1.83	1.75	1.78	1.84	2.05
	250	500	500	2.28	2.09	2.04	2.07	2.14	2.34
	500	1000	1000	2.29	2.13	2.11	2.16	2.19	2.36
Unbalanced	100	100	50	1.79	1.32	1.33	1.40	1.57	1.99
	200	200	100	1.99	1.44	1.36	1.51	1.71	2.14
	500	500	250	2.22	1.72	1.57	1.78	2.02	2.35
	1000	1000	500	2.23	1.75	1.61	1.76	1.98	2.36

Table 4.1 presents the results of the type I error rate simulations. For all scenarios, type I error rates are controlled as tests with a significance level of 2.5%. In particular, when $0 < a_{02} < 1$, they are more conservative than those when $a_{02} = 0$ or 1. A similar tendency was observed for the binomial distribution case Doi et al. (2017b).

Table 4.2: Power of the Bayesian method under $H_1 : \bar{\lambda}_1 < \bar{\lambda}_2 + \Delta$ ($\bar{\lambda}_1 = 0.3, \bar{\lambda}_2 = 0.3, \Delta = 0.1$).

				a_{02}					
Scenario	n_{02}	n_1	n_2	0.0	0.1	0.25	0.5	0.75	1.0
Balanced 1	100	100	100	16.91	15.95	16.58	18.07	19.01	20.70
	200	200	200	31.79	31.72	33.36	35.94	37.38	39.84
	500	500	500	68.57	70.04	72.68	75.78	77.46	79.09
	1000	1000	1000	93.78	94.70	95.92	97.01	97.49	97.79
Balanced 2	50	100	100	16.94	15.94	16.40	17.23	17.71	19.26
	100	200	200	32.16	31.59	32.60	34.11	35.17	37.03
	250	500	500	68.74	69.13	70.68	72.84	73.89	75.64
	500	1000	1000	93.85	94.30	95.05	95.92	96.38	96.74
Unbalanced	100	100	50	12.59	11.73	13.14	15.32	16.84	19.54
	200	200	100	23.70	23.48	26.11	30.32	33.48	36.76
	500	500	250	53.12	56.66	62.45	69.10	72.58	75.32
	1000	1000	500	82.89	87.23	91.62	94.70	96.04	96.68

Table 4.2 shows the results of the power simulations. When $0 < a_{02} < 0.25$ and n_{02}, n_1, n_2 are small, powers are less than those for $a_{02} = 0$. This may be related to the conservativeness of the type I error rates shown in Table 4.1. In contrast, when $a_{02} \geq 0.5$ or n_{02}, n_1, n_2 are large, powers increased monotonically with a_{02}. Therefore, with an adequate historical trial size or suitable amount of borrowed information, the powers of our Bayesian method can be improved by historical data.

4.5 Real data analysis

Massacesi et al. (2014) performed a multicenter, randomized, controlled, single-blinded, non-inferiority trial in which patients with relapsing-remitting multiple sclerosis received azathioprine (AZA) or interferon beta (IFN-β). We refer to this trial as the present trial. The primary objective of this trial was to demonstrate that the annualized relapse rate (RR) over two years of AZA was not inferior to that of IFN-β. Here, we consider two-sided 95% CI of RR instead of one-sided 95% CI in Massacesi et al. (2014) because two-sided 95% CI or one-sided 97.5% CI is more conventional in the pharmaceutical industry after ICH E9 (ICH E9 Expert Working Group (1999)) was issued. The results for the annualized relapse rate are presented in Table 4.3. The two-sided 95% CI of RR was [0.43, 1.03], demonstrating the non-inferiority under the non-inferiority margin of 1.23 with a significance level of 2.5%. Next, switching to superiority, it was not demonstrated with a significance level of 2.5% because the upper limit of the CI exceeded 1.00.

Table 4.3: Annualized relapse rates of the present trial Massacesi et al. (2014).

	AZA ($n = 62$)	IFN-β ($n = 65$)	Rate Ratio [two-sided 95% CI]
Total person-years	126	132	
Total number of relapses	33	52	
Annualized relapse rate	0.26	0.39	0.67 [0.43, 1.03]

Next, we apply our Bayesian method. We borrow information for the IFN-β group from the trial in Cohen et al. (2012), which is the trial of alemtuzumab vs interferon β-1a (IFN-β-1a). Cohen et al. (2012) and Massacesi et al. (2014) showed similar relapse rates for IFN-β group. We refer to this trial Cohen et al. (2012) as the historical trial.

Table 4.4: Annualized relapse rates of the historical trial Cohen et al. (2012).

	IFN-β-1a ($n = 187$)	Alemtuzumab ($n = 376$)
Total number of relapses	122	119
Annualized relapse rate	0.39	0.18

Table 4.4 presents the results of the annualized relapse rate of the historical trial Cohen et al. (2012). Because total person-years in each group was not explicitly stated, we estimated this information from the total number of relapses and annualized relapse rate.

We let λ_1 and λ_2 be the relapse rates of the AZA and IFN-β groups, respectively. First, we construct the prior of λ_2 based on the historical trial data. From Table 4.4, the total person-years of the IFN-β-1a group is estimated as $[122/0.39] = 312$. Next, we let the prior of λ_2 be $Ga([122a_{02}], 312a_{02})$. Because this historical trial does not include an AZA group, we let the prior distribution of AZA group be $\lambda_1 \propto 1$.

Table 4.5 shows the results with $a_{02} = 0, 0.1, 0.2, 0.3, 0.5$ and $\Delta = 0, 0.05, 0.10$. $\Delta > 0$ and $\Delta = 0$ indicates the non-inferiority and superiority test, respectively. Here, ESS is the prior

Table 4.5: Result of the MS data analysis.

		prior	posterior		Δ		
a_{02}	ESS	$Ga(\alpha_2, \beta_2)$	$Ga(a_1, b_1)$	$Ga(a_2, b_2)$	0	0.05	0.1
0	(0)	λ_2^{-1}	$Ga(34, 126.0)$	$Ga(52, 132.0)$	0.959	0.993	0.999
0.1	12	$Ga(12, 31.2)$	$Ga(34, 126.0)$	$Ga(64, 163.2)$	0.965	0.994	0.999
0.2	24	$Ga(24, 62.4)$	$Ga(34, 126.0)$	$Ga(76, 194.4)$	0.968	0.995	0.999
0.3	36	$Ga(36, 93.6)$	$Ga(34, 126.0)$	$Ga(88, 225.6)$	0.971	0.996	1.000
0.5	61	$Ga(61, 156.0)$	$Ga(34, 126.0)$	$Ga(113, 288.0)$	0.978	0.997	1.000

effective sample size defined by Morita et al. (2008), which characterize the amount of information contained in the prior distribution. In each a_{02} and Δ, non-inferiority is demonstrated. Furthermore, the superiority is also shown when $a_{02} \geq 0.5$. This shows that our Bayesian method may be more efficient than the frequentist method when we want to switch the objective from non-inferiority to superiority and sufficient quality and quantity of the historical data.

4.6 Conclusion

In this chapter, we considered the Bayesian framework of the non-inferiority test and switching to the superiority test. First, we derived the exact expression for the posterior probability of the alternative hypothesis being true, and provided a sample SAS program. Next, we gave the framework cooperating the historical information flexibly. After that, we showed the relationship between the Bayesian non-inferiority and superiority tests and that between the Bayesian non-inferiority probability and superiority conditional test. We can naturally apply our method to trials switching from non-inferiority to superiority.

From the Monte Carlo simulations, the type I error rates were controlled for all planned scenarios. Additionally, based on the real data analysis, our Bayesian method can improve the power with a suitable amount of borrowed information. However, when $0 < a_{02} < 1$, it tended to be too conservative. This property will need to be improved in the future work.

Appendix: Sample SAS code for Pr$(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2)$

When $\Delta = 1$ and the posterior distributions are $\lambda_1 \sim Beta(120, 100)$ and $\lambda_2 \sim Beta(50, 200)$, the probability Pr$(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2)$ can be calculated by the following code based on (4.1).

```
%let  a1 = 120;   %let  b1 = 100;
%let  a2 = 50;    %let  b2 = 200;
%let  delta = 1;

data d1;
  a1 = &a1;  b1=&b1;  a2=&a2;  b2=&b2;  delta=&delta;
  p_NB = b2/(b1+b2);
  p0 = 0;
  do  j = 0  to  a1 -1 ;
       p0 = p0 + pdf('Poisson', j , delta*b1)
              * cdf('NEGBINOMIAL', a1-1- j, p_NB, a2);
  end;
  prob_Bayes = 1- p0;
  drop  j  p0;
run;
```

Chapter 5

Bayesian superiority and equivalence hypothesis testings and the p-value of the F-test for the variance of normal distributions

5.1 Introduction

In biomedical studies, we encounter occasions to compare the variances in variables of interest across different conditions. Such occasions may be divided into two different situations. The first situation is when we mainly focus on comparing the variances. For example, test-retest variabilities (TRV) of visual acuity measurements are compared across different degrees of optical defocus in Rosser et al. (2004) and across different methods of scoring in Bosch and Wall (1997). The second situation is when we mainly focus on the location parameters (e.g., mean), and we want to check the assumption about the variances in the statistical method for comparing them. In many clinical trials with continuous outcomes, linear (mixed) models including t-test, ANOVA, and ANCOVA are used as the method of the primary analysis. Based on whether the variances are equal or not, we may change the statistical method (e.g., Student's t-test or Welch's t-test) because an inappropriate choice of the method may lead to incorrect conclusions. See, for example, Welch (1938) and

Glass et al. (1972). Therefore, to choose the correct method is important. However, for many clinical trials in this situation, the tests comparing the variances are known to have lower power than expected. See, for example, Markowski and Markowski (1990) and Wilcox (1995). This may occur because the sample sizes are calculated for comparing location parameters, which reduces the power of the test for comparing variances.

For a general two-group comparison of parameters, Bayesian approaches have gained increasing attention for their potential superiority in decision making compared to conventional frequentist methods, because a Bayesian approach can borrow strength from the historical data. For example, with a binomial distribution $B(n_i, p_i)$, Altham (1969), Kawasaki and Miyaoka (2012b), Zaslavsky (2013), and Kawasaki et al. (2014) considered the posterior probability Pr$(p_1 > p_1 \mid X_1, X_2)$. For the Poisson distribution $Po(\lambda_i)$, Kawasaki and Miyaoka (2012a) and Doi (2016) considered Pr$(\lambda_1 < \lambda_2 \mid X_1, X_2)$. Kawasaki and Miyaoka (2012b) referred to these types of probabilities as Bayesian indexes. For both distributions, the Bayesian indexes were shown to be expressed by the hypergeometric series, and the relationship between the Bayesian indexes and the p-values of conventional frequentist tests were investigated.

In this chapter, we consider the problem of comparing the variances of two normal populations. F-test is most frequently used in this situation. To achieve a more effective decision than possible with the F-test by borrowing strength from the historical data, we propose a Bayesian index of superiority and equivalence for comparing the variances of two groups of normally distributed data.

The remainder of this chapter is structured as follows. In Section 5.2, we propose Bayesian indexes of superiority for three situations, express these indexes by the hypergeometric series and the cumulative distribution functions of well-known distributions, and investigate their relationship with the p-values of the F-test. In Section 5.3, we propose the Bayesian index of equivalence, which is also expressed by the hypergeometric series and the cumulative distributions functions. In Section 5.4, we present the results of a Monte Carlo simulation to investigate the properties of $\theta_e(\Delta) \geq \gamma$ for several Δ and γ values used in the Bayesian index of equivalence. In Section 5.5, we apply the Bayesian indexes to analyses of real data from actual clinical trials. Finally, we offer concluding remarks and highlight the prospects of these indexes in Section 5.6.

5.2 Bayesian Index of Superiority

5.2.1 Definition and the Fundamental Theorem

For $i = 1, 2$, and $n_1, n_2 \in \mathbb{N}$, let $X_{i1}, \ldots, X_{in_i}$ be independent normal random variables with mean μ_i and variance σ_i^2. Let the realized values of $X_{i1}, \cdots, X_{in_i}$ be denoted by $\mathbf{x}_i = (x_{i1}, \ldots, x_{in_i})'$. For Bayesian analysis, let the prior distribution of σ_i^2 be the scaled inverse χ^2 distribution $Scaled$-inv-$\chi^2(\nu_i, \tau_i^2)$ for $\nu_i, \tau_i^2 > 0$, whose probability density function is

$$f(\sigma_i^2 \mid \nu_i, \tau_i^2) = \frac{\left(\nu_i \tau_i^2/2\right)^{\nu_i/2} \left(\sigma_i^2\right)^{-\nu_i/2-1}}{\Gamma\left(\nu_i/2\right)} \exp\left(-\frac{\nu_i \tau_i^2}{2\sigma_i^2}\right).$$

This is equivalent to the inverse gamma distribution Inv-$Ga\left(\nu_i/2, \nu_i\tau_i^2/2\right)$. To compare the variances of two groups, we propose the Bayesian index of superiority as follows:

$$\theta = \Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2).$$

In the following description, we first consider the case where μ_1 and μ_2 are known, and next consider the case where both means are unknown. In each case, the following theorem is crucial.

Theorem 5.1 (Doi et al. (2017a))**.** If the (marginal) posterior distribution of σ_i^2 is Inv-$Ga(a_i, b_i)$ for $i = 1, 2$, then the Bayesian index $\theta = \Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$ has the following three expressions:

$$\begin{aligned}\Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) =& 1 - \frac{1}{a_2 B(a_1, a_2)} \left(\frac{b_2}{b_1 + b_2}\right)^{a_2} \cdot {}_2F_1\left(a_2, 1 - a_1; 1 + a_2; \frac{b_2}{b_1 + b_2}\right) \\ =& I_{\frac{b_1}{b_1+b_2}}(a_1, a_2) \\ =& F_{2a_1, 2a_2}\left(\frac{b_1/a_1}{b_2/a_2}\right),\end{aligned}$$

where

$${}_2F_1(a, b; c; z) = \sum_{t=0}^{\infty} \frac{(a)_t (b)_t}{(c)_t} \cdot \frac{z^t}{t!} \quad (|z| < 1)$$

is the hypergeometric series, and $(k)_t = k(k + 1) \cdots (k + t - 1)$ for $t \in \mathbb{N}$ and $(k)_0 = 1$ is the

Pochhammer symbol,

$$F_{\nu_1,\nu_2}(x) = \int_0^x \frac{1}{zB\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)} \left(\frac{\nu_1 z}{\nu_1 z + \nu_2}\right)^{\frac{\nu_1}{2}} \left(\frac{\nu_2}{\nu_1 z + \nu_2}\right)^{\frac{\nu_2}{2}} dz$$

is the cumulative distribution function of the F distribution $F(\nu_1, \nu_2)$, and

$$I_x(a,b) = \frac{1}{B(a,b)} \int_0^x t^{a-1}(1-t)^{b-1} dt$$

is the cumulative distribution function of the beta distribution $Beta(a,b)$, also known as the regularized incomplete beta function.

Proof. Let $\lambda_i = 1/\sigma_i^2$ be the precision, then

$$\begin{aligned} \theta &= \Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) \\ &= \Pr(\lambda_1 < \lambda_2 \mid \mathbf{x}_1, \mathbf{x}_2). \end{aligned} \tag{5.1}$$

When the (marginal) posterior distribution of σ_i^2 is $Inv\text{-}Ga\,(a_i, b_i)$, the (marginal) posterior distribution of λ_i is $Ga(a_i, b_i)$, whose probability density function is $f(\lambda_i \mid a_i, b_i) = b_i^{a_i}/\Gamma(a_i) \cdot \lambda_i^{a_i-1}\exp(-b_i\lambda_i)$. Hence, (5.1) is the Bayesian index for the Poisson parameters defined in Kawasaki and Miyaoka (2012a). Therefore, Theorem 5.1 follows from Kawasaki and Miyaoka (2012a) and Theorem 1 of Doi (2016). □

From the cumulative distribution function expressions in Theorem 5.1, $\Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$ can be quite easily calculated using standard statistical software.

Remark 3. For Theorem 5.1, since only the (marginal) posterior distribution of σ_i^2 is supposed as the inverse gamma distribution, the prior distribution of σ_i^2 may be improper as long as the posterior is the inverse gamma.

5.2.2 Case 1: μ_1 and μ_2 are Known

5.2.2.1 Calculation of the Bayesian Index of Superiority

In this case, we denote the likelihood of σ_i^2 by $L(\sigma_i^2 \mid \mathbf{x}_i, \mu_i)$. Since we suppose that the prior distribution of σ_i^2 is $Scaled$-inv-$\chi^2(\nu_i, \tau_i^2)$, the posterior distribution of σ_i^2 can be derived as follows:

$$
\begin{aligned}
& f(\sigma_i^2 \mid \mathbf{x}_i, \mu_i, \nu_i, \tau_i^2) \\
\propto & L(\sigma_i^2 \mid \mathbf{x}_i, \mu_i) \cdot f(\sigma_i^2 \mid \nu_i, \tau_i^2) \\
= & \frac{1}{(2\pi\sigma_i^2)^{n_i/2}} \exp\left(-\sum_{j=1}^{n_i} \frac{(x_{ij}-\mu_i)^2}{2\sigma_i^2}\right) \cdot \frac{(\nu_i\tau_i^2/2)^{\nu_i/2} (\sigma_i^2)^{-\nu_i/2-1}}{\Gamma(\nu_i/2)} \exp\left(-\frac{\nu_i\tau_i^2}{2\sigma_i^2}\right) \\
\propto & \left(\sigma_i^2\right)^{-(\nu_i+n_i)/2-1} \exp\left(-\frac{\nu_i\tau_i^2 + n_i \cdot T_i^2}{2\sigma_i^2}\right),
\end{aligned}
$$

where

$$T_i^2 = \frac{1}{n_i}\sum_{j=1}^{n_i}(x_{ij}-\mu_i)^2.$$

Hence, the posterior distribution of σ_i^2 is Inv-$Ga((\nu_i+n_i)/2, (\nu_i\tau_i^2 + n_i \cdot T_i^2)/2)$. Therefore, from Theorem 5.1, we have

$$\Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = F_{\nu_1+n_1,\nu_2+n_2}\left(\frac{(\nu_1\tau_1^2 + n_1 \cdot T_1^2)/(\nu_1+n_1)}{(\nu_2\tau_2^2 + n_2 \cdot T_2^2)/(\nu_2+n_2)}\right). \tag{5.2}$$

5.2.2.2 The Relationship between the Bayesian Index of Superiority and the p-value of the one-sided F-test

Here, we consider the F-test with $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_1 : \sigma_1^2 > \sigma_2^2$ when μ_1 and μ_2 are known. Under H_0, the test statistics T_1^2/T_2^2 follow $F(n_1, n_2)$. Hence, the p-value is calculated as

$$
\begin{aligned}
p &= \int_{T_1^2/T_2^2}^{\infty} \frac{1}{zB\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \left(\frac{n_1 z}{n_1 z + n_2}\right)^{\frac{n_1}{2}} \left(\frac{n_2}{n_1 z + n_2}\right)^{\frac{n_2}{2}} dz \\
&= 1 - F_{n_1,n_2}(T_1^2/T_2^2).
\end{aligned} \tag{5.3}
$$

Then, the following theorem holds.

Theorem 5.2 (Doi et al. (2017a))**.** If μ_1 and μ_2 are known and the prior distribution of σ_i^2 is $Scaled\text{-}inv\text{-}\chi^2(\nu_i, \tau_i^2)$ for $i = 1, 2$, then the following relation holds between the Bayesian index $\theta = \Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$ and the one-sided p-value of the F-test with $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_1 : \sigma_1^2 > \sigma_2^2$:

$$\lim_{(\nu_1,\nu_2)\to(0,0)} \Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = 1 - p.$$

Proof. From (5.2) and (5.3),

$$\lim_{(\nu_1,\nu_2)\to(0,0)} \Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = F_{n_1,n_2}(T_1^2/T_2^2) = 1 - p$$

holds. □

Remark 4. For the prior distribution,

$$\begin{aligned} f(\sigma_i^2 \mid \nu_i, \tau_i^2) &= \frac{(\nu_i\tau_i^2/2)^{\nu_i/2}\,(\sigma_i^2)^{-\nu_i/2-1}}{\Gamma(\nu_i/2)} \exp\left(-\frac{\nu_i\tau_i^2}{2\sigma_i^2}\right) \\ &\propto (\sigma_i^2)^{-\nu_i/2-1} \exp\left(-\frac{\nu_i\tau_i^2}{2\sigma_i^2}\right) \\ &\xrightarrow[\nu_i\to 0]{} (\sigma_i^2)^{-1}, \end{aligned}$$

when the prior distribution of σ_i^2 is $f(\sigma_i^2) \propto (\sigma_i^2)^{-1}$, which is improper, the posterior distribution of σ_i^2 is $Inv\text{-}Ga(n_i/2, n_i \cdot T_i^2/2)$. Therefore, as stated in Remark 3, the Bayesian index can be expressed by Theorem 5.1 as

$$\Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = F_{n_1,n_2}\left(T_1^2/T_2^2\right).$$

Then, the following theorem holds.

Theorem 5.3 (Doi et al. (2017a))**.** If μ_1 and μ_2 are known and the prior distribution of σ_i^2 is $f(\sigma_i^2) \propto (\sigma_i^2)^{-1}$ for $i = 1, 2$, then

$$\Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = 1 - p$$

holds.

Since ν_i is the prior effective sample size of $Scaled\text{-}inv\text{-}\chi^2(\nu_i, \tau_i^2)$ as defined in Morita et al. (2008), Theorem 5.3 can be interpreted as follows: the Bayesian index with prior effective sample size 0 for both groups is equal to $(1-p)$ of the one-sided F-test. Furthermore, with the prior $Scaled\text{-}inv\text{-}\chi^2(\nu_i, \tau_i^2)$, the Bayesian index can be interpreted as equal to the F-test with the prior information of α_i additional samples.

5.2.3 Case 2: μ_1 and μ_2 are Unknown

In this case, we consider two types of the prior distributions of (μ_i, σ_i^2). In each type, we denote the likelihood of (μ_i, σ_i^2) by $L(\mu_i, \sigma_i^2 \mid \mathbf{x}_i)$. Furthermore, in the following, we denote the probability density function of the normal inverse gamma distribution $NIG(\mu_0, k, \alpha, \beta)$ for $\mu_0 \in \mathbb{R}, k, \alpha, \beta > 0$ by

$$f(\mu, \sigma^2 \mid \mu_0, k, \alpha, \beta) = \sqrt{\frac{k}{2\pi\sigma^2}} \exp\left(-\frac{k(\mu-\mu_0)^2}{2\sigma^2}\right) \times \frac{\beta^\alpha}{\Gamma(\alpha)}(\sigma^2)^{-\alpha-1} \exp\left(-\frac{\beta}{\sigma^2}\right).$$

When (μ, σ^2) follows $NIG(\mu_0, k, \alpha, \beta)$, the marginal distribution of σ^2 is $Inv\text{-}Ga(\alpha, \beta)$.

5.2.3.1 Calculation of the Bayesian index of superiority for the scaled inverse χ^2 variance prior

We first suppose that the prior distribution of μ_i is non-informative, i.e., $f(\mu_i) \propto 1$, and the prior distribution of σ_i^2 is $Scaled\text{-}inv\text{-}\chi^2(\nu_i, \tau_i^2)$. Then, the prior distribution of (μ_i, σ_i^2) is $Scaled\text{-}inv\text{-}\chi^2(\nu_i, \tau_i^2)$. Here, the posterior distribution of (μ_i, σ_i^2) can be derived as

$$\begin{aligned}
& f(\mu_i, \sigma_i^2 \mid \mathbf{x}_i, \nu_i, \tau_i^2) \\
\propto & L(\mu_i, \sigma_i^2 \mid \mathbf{x}_i) \cdot f(\mu_i, \sigma_i^2 \mid \nu_i, \tau_i^2) \\
= & \frac{1}{(2\pi\sigma_i^2)^{n_i/2}} \exp\left(-\sum_{j=1}^{n_i} \frac{(x_{ij}-\mu_i)^2}{2\sigma_i^2}\right) \cdot \frac{(\nu_i\tau_i^2/2)^{\nu_i/2}\,(\sigma_i^2)^{-\nu_i/2-1}}{\Gamma(\nu_i/2)} \exp\left(-\frac{\nu_i\tau_i^2}{2\sigma_i^2}\right) \\
\propto & \left(\sigma_i^2\right)^{-1/2} \exp\left(-\frac{n_i(\mu_i-\bar{x}_i)^2}{2\sigma_i^2}\right) \cdot \left(\sigma_i^2\right)^{-\frac{\nu_i+n_i-1}{2}-1} \exp\left(-\frac{\nu_i\tau_i^2 + (n_i-1)\cdot S_i^2}{2\sigma_i^2}\right),
\end{aligned}$$

where

$$\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}, S_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2.$$

Therefore, the posterior distribution of (μ_i, σ_i^2) is $NIG\,(\bar{x}_i, n_i, (\nu_i + n_i - 1)/2, (\nu_i\tau_i^2 + (n_i - 1)\cdot S_i^2)/2)$, and the marginal posterior distribution of σ_i^2 is Inv-$Ga((\nu_i + n_i - 1)/2, (\nu_i\tau_i^2 + (n_i - 1)\cdot S_i^2)/2)$. Then, the Bayesian index can be expressed by Theorem 5.1 as

$$\Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = F_{\nu_1+n_1-1,\nu_2+n_2-1}\left(\frac{(\nu_1\tau_1^2 + (n_1 - 1)\cdot S_1^2)/(\nu_1 + n_1 - 1)}{(\nu_2\tau_2^2 + (n_2 - 1)\cdot S_2^2)/(\nu_2 + n_2 - 1)}\right). \tag{5.4}$$

5.2.3.2 The relationship between the Bayesian index of superiority for the scaled inverse χ^2 variance prior and the p-value of the one-sided F-test

Here, we consider the F-test with $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_1 : \sigma_1^2 > \sigma_2^2$ when μ_1 and μ_2 are unknown. Under H_0, the test statistics S_1^2/S_2^2 follow $F(n_1 - 1, n_2 - 1)$. Therefore, the p-value is calculated as

$$\begin{aligned} p =& \int_{S_1^2/S_2^2}^{\infty} \frac{1}{zB\left(\dfrac{n_1 - 1}{2}, \dfrac{n_2 - 1}{2}\right)} \\ &\times \left(\frac{(n_1 - 1)z}{(n_1 - 1)z + (n_2 - 1)}\right)^{\dfrac{n_1 - 1}{2}} \left(\frac{n_2 - 1}{(n_1 - 1)z + (n_2 - 1)}\right)^{\dfrac{n_2 - 1}{2}} dz \\ =& 1 - F_{n_1-1,n_2-1}(S_1^2/S_2^2). \end{aligned} \tag{5.5}$$

Then, the following theorem holds.

Theorem 5.4 (Doi et al. (2017a))**.** If the prior distribution of μ_i is non-informative, i.e., $f(\mu_i) \propto 1$, and that of σ_i^2 is $Scaled$-inv-$\chi^2(\nu_i, \tau_i^2)$ for $i = 1, 2$, respectively, then the following relation holds between the Bayesian index $\theta = \Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$ and the one-sided p-value of the F-test with $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_1 : \sigma_1^2 > \sigma_2^2$:

$$\lim_{(\nu_1,\nu_2)\to(0,0)} \Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = 1 - p.$$

Proof. From (5.4) and (5.5),

$$\lim_{(\nu_1,\nu_2)\to(0,0)} \Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = F_{n_1-1,n_2-1}(S_1^2/S_2^2) = 1 - p$$

holds. □

Remark 5. For the prior distribution,

$$\begin{aligned} f(\mu_i, \sigma_i^2 \mid \nu_i, \tau_i^2) &= \frac{(\nu_i \tau_i^2/2)^{\nu_i/2} (\sigma_i^2)^{-\nu_i/2-1}}{\Gamma(\nu_i/2)} \exp\left(-\frac{\nu_i\tau_i^2}{2\sigma_i^2}\right) \\ &\propto (\sigma_i^2)^{-\nu_i/2-1} \exp\left(-\frac{\nu_i\tau_i^2}{2\sigma_i^2}\right) \\ &\xrightarrow[\nu_i\to 0]{} (\sigma_i^2)^{-1}, \end{aligned}$$

when the prior distribution of (μ_i, σ_i^2) is $f(\mu_i, \sigma_i^2) \propto (\sigma_i^2)^{-1}$, the posterior distribution of (μ_i, σ_i^2) is $NIG(\bar{x}_i, n_i, (n_i-1)/2, (n_i-1)\cdot S_i^2/2)$ and the marginal posterior distribution of σ_i^2 is Inv-$Ga((n_i-1)/2, (n_i-1)\cdot S_i^2/2)$. Therefore, the Bayesian index can be expressed by Theorem 5.1 as

$$\Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = F_{n_1-1,n_2-2}\left(S_1^2/S_2^2\right).$$

Then, the following theorem holds.

Theorem 5.5 (Doi et al. (2017a))**.** If the prior distribution of (μ_i, σ_i^2) is $f(\mu_i, \sigma_i^2) \propto (\sigma_i^2)^{-1}$ for $i = 1, 2$, then

$$\Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = 1 - p$$

holds.

5.2.3.3 Calculation of the Bayesian index of superiority for the normal inverse gamma prior

We next suppose that the prior distribution is $\mu_i \mid \sigma_i^2 \sim N(\mu_{0,i}, \sigma_i^2/k_i)$ and $\sigma_i^2 \sim Scaled$-inv-$\chi^2(\nu_i, \tau_i^2) = Inv$-$Ga(\nu_i/2, \nu_i\tau_i^2/2)$. Then, the prior distribution of (μ_i, σ_i^2) is the normal inverse gamma distribution $NIG\left(\mu_{0,i}, k_i, \nu_i/2, \nu_i\tau_i^2/2\right)$. Hence, the posterior distribution of (μ_i, σ_i^2) can be derived as

$$\begin{aligned}
& f(\mu_i, \sigma_i^2 \mid \mathbf{x}_i, \mu_{0,i}, k_i, \nu_i, \tau_i^2) \\
\propto & L(\mu_i, \sigma_i^2 \mid \mathbf{x}_i) \cdot f(\mu_i, \sigma_i^2 \mid \mu_{0,i}, k_i, \nu_i, \tau_i^2) \\
= & \frac{1}{(2\pi\sigma_i^2)^{n_i/2}} \exp\left(-\sum_{j=1}^{n_i} \frac{(x_{ij}-\mu_i)^2}{2\sigma_i^2}\right) \cdot \left(\frac{k_i}{2\pi\sigma_i^2}\right)^{1/2} \exp\left(-\frac{k_i(\mu_i-\mu_{0,i})^2}{2\sigma_i^2}\right) \\
& \times \frac{\left(\nu_i\tau_i^2/2\right)^{\nu_i/2} \left(\sigma_i^2\right)^{-\nu_i/2-1}}{\Gamma\left(\nu_i/2\right)} \exp\left(-\frac{\nu_i\tau_i^2}{2\sigma_i^2}\right) \\
\propto & \left(\sigma_i^2\right)^{-1/2} \exp\left(-\frac{(k_i+n_i)(\mu_i-\mu_{ni})^2}{2\sigma_i^2}\right) \\
& \times \left(\sigma_i^2\right)^{-\frac{\nu_i+n_i}{2}-1} \cdot \exp\left(-\frac{\nu_i\tau_i^2 + (n_i-1)\cdot S_i^2 + \dfrac{k_i n_i(\mu_{0,i}-\bar{x}_i)^2}{k_i+n_i}}{2\sigma_i^2}\right),
\end{aligned}$$

where

$$\begin{aligned}
\mu_{ni} &= \frac{k_i\mu_{0,i} + n_i\bar{x}_i}{k_i+n_i}, k_{ni} = k_i + n_i, \\
a_{ni} &= \frac{\nu_i+n_i}{2}, b_{ni} = \frac{\nu_i\tau_i^2 + (n_i-1)\cdot S_i^2}{2} + \frac{k_i n_i(\mu_{0,i}-\bar{x}_i)^2}{2(k_i+n_i)}.
\end{aligned}$$

Then, the posterior distribution of (μ_i, σ_i^2) is $NIG\left(\mu_{ni}, k_{ni}, a_{ni}, b_{ni}\right)$. Hence, the marginal posterior distribution of σ_i^2 is $Inv\text{-}Ga(a_{ni}, b_{ni})$. Therefore, the Bayesian index can be expressed from Theorem 5.1 as

$$\theta = F_{2a_{n1},2a_{n2}}\left(\frac{\dfrac{\nu_1\tau_1^2 + (n_1-1)\cdot S_1^2}{\nu_1+n_1} + \dfrac{k_1 n_1(\mu_{0,1}-\bar{x}_1)^2}{(\nu_1+n_1)(k_1+n_1)}}{\dfrac{\nu_2\tau_2^2 + (n_2-1)\cdot S_2^2}{\nu_2+n_2} + \dfrac{k_2 n_2(\mu_{0,2}-\bar{x}_2)^2}{(\nu_2+n_2)(k_2+n_2)}}\right). \tag{5.6}$$

5.2.3.4 The relationship between the Bayesian index of superiority for the normal inverse gamma prior and the p-value of the one-sided F-test

Then, the following theorem holds.

Theorem 5.6 (Doi et al. (2017a))**.** If the prior distribution of (μ_i, σ_i^2) is $NIG(\mu_{0,i}, k_i, \nu_i/2, \nu_i\tau_i^2/2)$

for $i = 1, 2$, then the following relation holds between the Bayesian index $\theta = \Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$ and the one-sided p-value of the F-test with $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_1 : \sigma_1^2 > \sigma_2^2$:

$$\lim_{(\nu_1, k_1, \tau_1^2, \nu_2, k_2, \tau_2^2) \to (-1,0,0,-1,0,0)} \Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = 1 - p.$$

Proof. From (5.6) and (5.5),

$$\begin{aligned} \lim_{(\nu_1, k_1, \tau_1^2, \nu_2, k_2, \tau_2^2) \to (-1,0,0,-1,0,0)} \Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) &= F_{n_1-1, n_2-1}(S_1^2/S_2^2) \\ &= 1 - p \end{aligned}$$

holds. □

Remark 6. For the prior distribution,

$$\begin{aligned} &f(\mu_i, \sigma_i^2 \mid \mu_{0,i}, k_i, \nu_i, \tau_i^2) \\ &= \sqrt{\frac{k_i}{2\pi\sigma_i^2}} \exp\left(-\frac{k_i(\mu_i - \mu_{0,i})^2}{2\sigma_i^2}\right) \cdot \frac{(\nu_i \tau_i^2/2)^{\nu_i/2}}{\Gamma(\nu_i/2)} (\sigma_i^2)^{-\nu_i/2-1} \exp\left(-\frac{\nu_i \tau_i^2}{2\sigma_i^2}\right) \\ &\propto (\sigma_i^2)^{-(\nu_i+1)/2-1} \exp\left(-\frac{\nu_i \tau_i^2 + k_i(\mu_i - \mu_{0,i})^2}{2\sigma_i^2}\right) \\ &\xrightarrow[(\nu_i, \tau_i^2, k_i) \to (-1,0,0)]{} (\sigma_i^2)^{-1}. \end{aligned}$$

As already shown in Theorem 5.5, if the prior distribution of (μ_i, σ_i^2) is $f(\mu_i, \sigma_i^2) \propto (\sigma_i^2)^{-1}$, then

$$\Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = 1 - p$$

holds.

5.2.4 Remark on the Prior Distribution

To utilize the historical data effectively, we here consider how to construct the prior distribution of (μ_i, σ_i^2). For $i = 1, 2$ and $j = 1, \cdots, n_{0,i}$, let the historical data $x_{0,ij}$ independently follow

$N(\mu_i, \sigma_i^2)$, and $\mathbf{x}_{0,i} = (x_{0,i1}, \ldots, x_{0,in_{0,i}})'$, and let

$$f_0(\mu_i, \sigma_i^2) \propto \left(\sigma_i^2\right)^{-1}.$$

Here, for $0 \leq \alpha_i \leq 1$, an example of the conditional power prior distribution, defined in Ibrahim and Chen (2000), is

$$\begin{aligned} f(\mu_i, \sigma_i^2) \propto & L(\mu_i, \sigma_i^2 \mid \mathbf{x}_{0,i})^{\alpha_i} \cdot f_0(\mu_i, \sigma_i^2) \\ \propto & \left\{ \frac{1}{(2\pi\sigma_i^2)^{n_{0,i}/2}} \exp\left(-\frac{1}{2\sigma_i^2} \sum_{j=1}^{n_{0,i}} (x_{0,ij} - \mu_i)^2 \right) \right\}^{\alpha_i} \cdot \left(\sigma_i^2\right)^{-1} \\ = & \left(\sigma_i^2\right)^{-1/2} \exp\left(-\frac{\alpha_i n_{0,i}(\mu_i - \bar{x}_{0,i})^2}{2\sigma_i^2} \right) \\ & \times \left(\sigma_i^2\right)^{-(\alpha_i n_{0,i}-1)/2-1} \exp\left(-\frac{\alpha_i (n_{0,i}-1) \cdot S_{0,i}^2}{2\sigma_i^2} \right), \end{aligned} \tag{5.7}$$

where

$$\bar{x}_{0,i} = \frac{1}{n_{0,i}} \sum_{j=1}^{n_{0,i}} x_{0,ij}, S_{0,i}^2 = \frac{1}{n_{0,i}-1} \sum_{j=1}^{n_{0,i}} (x_{0,ij} - \bar{x}_{0,i})^2.$$

Then, the prior distribution of (μ_i, σ_i^2) is the normal inverse gamma distribution $NIG(\bar{x}_{0,i}, \alpha_i n_{0,i}, (\alpha_i n_{0,i} - 1)/2, \alpha_i(n_{0,i} - 1) \cdot S_{0,i}^2/2)$ when $\alpha_i n_{0,i} > 1$. Hence, the marginal prior distribution of σ_i^2 is $Inv\text{-}Ga((\alpha_i n_{0,i} - 1)/2, \alpha_i(n_{0,i} - 1) \cdot S_{0,i}^2/2)$ when $\alpha_i n_{0,i} > 1$. In this situation, the next corollary directly follows from Theorem 5.6.

Corollary 5.7 (Doi et al. (2017a))**.** If the prior distribution of (μ_i, σ_i^2) is the conditional power prior described above for $i = 1, 2$, then

$$\lim_{(\alpha_1,\alpha_2)\to(0,0)} \Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = 1 - p$$

holds.

5.3 Bayesian Index of Equivalence

Next, we propose the Bayesian index of equivalence for Δ satisfying $1 < \Delta$ as follows:

$$\theta_e(\Delta) = \Pr(\, 1/\Delta < \sigma_1/\sigma_2 < \Delta \mid \mathbf{x}_1, \mathbf{x}_2) \,.$$

Here, we compare Δ and $1/\Delta$ not to the ratio of the variances but rather to the ratio of the standard deviations. Then, the following theorem holds.

Theorem 5.8 (Doi et al. (2017a))**.** If the (marginal) posterior distribution of σ_i^2 is Inv-$Ga(a_i, b_i)$ for $i = 1, 2$, then

$$\begin{aligned}
\Pr(\, 1/\Delta < \sigma_1/\sigma_2 < \Delta \mid \mathbf{x}_1, \mathbf{x}_2) =& F_{2a_1,2a_2}\left(\frac{b_1/a_1}{b_2/a_2} \cdot \Delta^2\right) - F_{2a_1,2a_2}\left(\frac{b_1/a_1}{b_2/a_2} \cdot \frac{1}{\Delta^2}\right) \\
=& I_{\frac{b_1 \cdot \Delta^2}{b_1 \cdot \Delta^2 + b_2}}(a_1, a_2) - I_{\frac{b_1/\Delta^2}{b_1/\Delta^2 + b_2}}(a_1, a_2) \\
=& \frac{1}{a_2 B(a_1, a_2)}\left(\frac{b_2}{b_1/\Delta^2 + b_2}\right)^{a_2} \cdot {}_2F_1\left(a_2, 1 - a_1; 1 + a_2; \frac{b_2}{b_1/\Delta^2 + b_2}\right) \\
& - \frac{1}{a_2 B(a_1, a_2)}\left(\frac{b_2}{b_1 \cdot \Delta^2 + b_2}\right)^{a_2} \cdot {}_2F_1\left(a_2, 1 - a_1; 1 + a_2; \frac{b_2}{b_1 \cdot \Delta^2 + b_2}\right).
\end{aligned}$$

Proof. Since $\Pr(\sigma_1/\sigma_2 = \Delta) = 0$,

$$\begin{aligned}
\Pr(\, 1/\Delta < \sigma_1/\sigma_2 < \Delta \mid \mathbf{x}_1, \mathbf{x}_2) &= \Pr(\sigma_1/\sigma_2 < \Delta \mid \mathbf{x}_1, \mathbf{x}_2) - \Pr(\sigma_1/\sigma_2 < 1/\Delta \mid \mathbf{x}_1, \mathbf{x}_2) \\
&= \Pr(\sigma_1^2/\sigma_2^2 < \Delta^2 \mid \mathbf{x}_1, \mathbf{x}_2) - \Pr(\sigma_1^2/\sigma_2^2 < 1/\Delta^2 \mid \mathbf{x}_1, \mathbf{x}_2).
\end{aligned}$$

Then, consider the posterior distribution of $\sigma_1^2/\sigma_2^2 = \lambda_2/\lambda_1$, where $\lambda_i = 1/\sigma_i^2$ is the precision for $i = 1, 2$. From Theorem 3 in Doi (2016) or (2.10) in Price and Bonett (2000),

$$\Pr(\lambda_2/\lambda_1 < c \mid \mathbf{x}_1, \mathbf{x}_2) = F_{2a_2,2a_1}\left(\frac{b_2/a_2}{b_1/a_1} \cdot c\right),$$

therefore

$$\begin{aligned}
\Pr(\, 1/\Delta < \sigma_1/\sigma_2 < \Delta \mid \mathbf{x}_1, \mathbf{x}_2) =& \Pr(\lambda_2/\lambda_1 < \Delta^2 \mid \mathbf{x}_1, \mathbf{x}_2) - \Pr(\lambda_2/\lambda_1 < 1/\Delta^2 \mid \mathbf{x}_1, \mathbf{x}_2) \\
=& F_{2a_2,2a_1}\left(\frac{b_2/a_2}{b_1/a_1} \cdot \Delta^2\right) - F_{2a_2,2a_1}\left(\frac{b_2/a_2}{b_1/a_1} \cdot \frac{1}{\Delta^2}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \left\{1 - F_{2a_1,2a_2}\left(\frac{b_1/a_1}{b_2/a_2} \cdot \frac{1}{\Delta^2}\right)\right\} - \left\{1 - F_{2a_1,2a_2}\left(\frac{b_1/a_1}{b_2/a_2} \cdot \Delta^2\right)\right\} \\
&\quad (\because F_{m,n}(1/x) = 1 - F_{n,m}(x)) \\
&= F_{2a_1,2a_2}\left(\frac{b_1/a_1}{b_2/a_2} \cdot \Delta^2\right) - F_{2a_1,2a_2}\left(\frac{b_1/a_1}{b_2/a_2} \cdot \frac{1}{\Delta^2}\right).
\end{aligned}
$$

The rest of the proof follows from Theorem 3 in Doi (2016). □

5.4 Simulation

We conducted a Monte Carlo simulation to investigate the property of "$\theta_e(\Delta) \geq \gamma$" for several values of Δ and γ. We used the conditional power prior distribution. The historical data $x_{0,ij}$ independently follow $N(0, \sigma_i^2)$ for $i = 1, 2; j = 1, \ldots, n_{0,i}$, and we consider $\alpha_1 = \alpha_2 = 1$. The present data x_{ij} independently follow $N(0, \sigma_i^2)$ for $i = 1, 2; j = 1, \ldots, n_i$. Here, let $n = n_1 = n_2 = 25, 50, 100, 200$, and $n_0 = n_{0,1} = n_{0,2} = 0, 25, 50, 100, 200$, with $n_{0,i} \leq n_i$ for $i = 1, 2$. Further, we take $\gamma = 0.90, 0.95$ and $\Delta = 1.10, 1.25, 1.50, 2.00$. We conducted 100,000 iterations for each scenario. For the first scenario, we set $\sigma_1 = \sigma_2 = 10$; that is, the variances are equal. As shown in Table 5.1, the percentage satisfying this condition heavily depended on the sample size. For the second scenario, we set $\sigma_1 = 15$ and $\sigma_2 = 10$; that is, $\sigma_1/\sigma_2 = 1.5$, so that the variances of group 1 are greater than those of group 2. As shown in Table 5.2, the percentages satisfying $\theta_e(1.50) \geq 0.90$ and $\theta_e(1.50) \geq 0.95$ show minimal dependence on the sample size when $n \geq 50$.

These results suggest that the decision of suitable values of Δ and γ must be considered depending on the situation. For example, if $n = 50$ and $n_0 \geq 25$, then $\theta_e(1.50) \geq 0.90$ or $\theta_e(1.50) \geq 0.95$ may be appropriate.

5.5 Application

The application of the Bayesian indexes of superiority and equivalence was evaluated using data from actual clinical trials, as shown in Table 5.3. Trial (a) and (b) are two selected trials shown in Table 1 of Gould (1991). Here, we supposed that trial (a) is a previous trial and trial (b) is the present trial, and $i = 1, 2$ indicate the placebo and drug A group, respectively. Therefore, we

Table 5.1: The percentage satisfying $\theta_e(\Delta) = \Pr(1/\Delta < \sigma_1/\sigma_2 < \Delta \mid \mathbf{x}_1, \mathbf{x}_2) \geq \gamma$ when $\sigma_1 = \sigma_2 = 10$.

n_0	n	$\gamma = 0.90$				$\gamma = 0.95$			
		$\theta_e(1.10)$	$\theta_e(1.25)$	$\theta_e(1.50)$	$\theta_e(2.00)$	$\theta_e(1.10)$	$\theta_e(1.25)$	$\theta_e(1.50)$	$\theta_e(2.00)$
0	25	0.00	0.00	48.03	95.92	0.00	0.00	0.00	90.87
25	25	0.00	0.00	87.31	99.94	0.00	0.00	76.07	99.82
0	50	0.00	0.00	87.53	99.95	0.00	0.00	75.81	99.81
25	50	0.00	45.08	97.06	100.00	0.00	0.00	93.19	100.00
50	50	0.00	64.46	99.31	100.00	0.00	41.04	98.16	100.00
0	100	0.00	64.56	99.37	100.00	0.00	49.84	98.23	100.00
25	100	0.00	76.76	99.88	100.00	0.00	59.48	99.57	100.00
50	100	0.00	84.82	99.97	100.00	0.00	71.66	99.89	100.00
100	100	0.00	93.67	100.00	100.00	0.00	86.49	100.00	100.00
0	200	0.00	93.70	100.00	100.00	0.00	86.48	100.00	100.00
25	200	0.00	95.86	100.00	100.00	0.00	90.93	100.00	100.00
50	200	0.00	97.50	100.00	100.00	0.00	93.95	100.00	100.00
100	200	5.09	98.93	100.00	100.00	0.00	97.26	100.00	100.00
200	200	43.87	99.84	100.00	100.00	0.00	99.46	100.00	100.00

utilized the data of trial (a) to specify the conditional power prior. We suppose that $\alpha = \alpha_1 = \alpha_2$, and take $\alpha = 0.0, 0.2, 0.5, 0.8, 1.0$. The prior distributions of (μ_i, σ_i^2) were derived from (5.7) with the following data

- $n_{0,1} = 47, \bar{x}_{0,1} = 3.04, S_{0,1}^2 = 9.20^2 \fallingdotseq 84.64$
- $n_{0,2} = 44, \bar{x}_{0,2} = 8.43, S_{0,2}^2 = 8.17^2 \fallingdotseq 66.75$,

and are shown in Table 5.4 for each α. Next, using the following data of trial (b)

- $n_1 = 53, \bar{x}_1 = 3.75, S_1^2 = 7.07^2 \fallingdotseq 49.98$
- $n_2 = 54, \bar{x}_2 = 10.20, S_2^2 = 9.39^2 \fallingdotseq 88.17$,

we derived the posterior distributions. The posterior distributions of (μ_i, σ_i^2) and the marginal posterior distributions of $\sigma_i^2$are shown in Table 5.5 and Table 5.6, respectively.

Finally, the Bayesian indexes are shown in Table 5.7.

For trial (a), the placebo group $(i = 1)$ showed a larger standard deviation than the drug A group $(i = 2)$. By contrast, for trial (b), the drug A group showed a larger standard deviation. According to the present data (trial (b)) only, that is, when $\alpha = 0$, θ is quite small, which makes

Table 5.2: The percentage satisfying $\theta_e(\Delta) = \Pr(1/\Delta < \sigma_1/\sigma_2 < \Delta \mid \mathbf{x}_1, \mathbf{x}_2) \geq \gamma$ when $\sigma_1 = 15$, and $\sigma_2 = 10$.

n_0	n	$\gamma = 0.90$				$\gamma = 0.95$			
		$\theta_e(1.10)$	$\theta_e(1.25)$	$\theta_e(1.50)$	$\theta_e(2.00)$	$\theta_e(1.10)$	$\theta_e(1.25)$	$\theta_e(1.50)$	$\theta_e(2.00)$
0	25	0.00	0.00	8.87	54.19	0.00	0.00	0.00	39.33
25	25	0.00	0.00	10.62	76.71	0.00	0.00	5.48	64.46
0	50	0.00	0.00	10.06	76.21	0.00	0.00	5.04	63.49
25	50	0.00	0.27	10.27	88.19	0.00	0.00	5.19	79.53
50	50	0.00	0.12	10.16	94.12	0.00	0.04	5.16	88.61
0	100	0.00	0.11	10.00	94.23	0.00	0.02	4.94	88.67
25	100	0.00	0.07	10.06	97.17	0.00	0.02	5.06	93.89
50	100	0.00	0.03	10.00	98.63	0.00	0.01	5.09	96.77
100	100	0.00	0.00	10.00	99.72	0.00	0.00	5.01	99.21
0	200	0.00	0.01	10.19	99.73	0.00	0.00	5.16	99.21
25	200	0.00	0.00	10.10	99.86	0.00	0.00	5.04	99.58
50	200	0.00	0.00	9.99	99.95	0.00	0.00	5.09	99.79
100	200	0.00	0.00	9.97	99.99	0.00	0.00	5.02	99.95
200	200	0.00	0.00	9.95	100.00	0.00	0.00	4.98	100.00

Table 5.3: Hypertention data in Gould (1991).

	Placebo $(i = 1)$			Drug A $(i = 2)$		
Trial	n	mean	SD	n	mean	SD
(a)	47	3.04	9.20	44	8.43	8.17
(b)	53	3.75	7.07	54	10.20	9.39

the variance of the placebo group seem greater. However, as α increases, i.e., the weight of the information of trial (a) increases, θ increases monotonically, and is no longer small. Furthermore, the p-value of the F-test with $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_1 : \sigma_1^2 > \sigma_2^2$ is 0.979, and, as shown in remark 6, is equal to $1 - \theta$ with $\alpha = 0.0$. Next, we consider the situation of $\theta_e(\Delta)$. Based on the result of the simulation in section 5.4, we assume that $\theta_e(1.50) \geq 0.95$ shows the equivalence, because the sample size is about 50 for both groups and for both the historical and present data. Then, when using only the present data ($\alpha = 0.0$) and $\alpha = 0.2$, the equivalence is not shown. By contrast, when $\alpha = 0.5, 0.8, 1.0$, that is, when the weight of the historical data is moderate to large, the equivalence is shown.

In order to apply these indexes to the real clinical trials, we have to consider whether α, Δ and γ can be pre-specified based on sufficiently reliable information. If we can pre-specify them suitably, we can determine the statistical method for comparing the means based on whether $\theta_e(\Delta) \geq \gamma$

Table 5.4: Prior distributions of (μ_i, σ_i^2).

α	Placebo $(i = 1)$	Drug A $(i = 2)$
0.0	$f(\mu_1, \sigma_1^2) \propto (\sigma_1^2)^{-1}$	$f(\mu_2, \sigma_2^2) \propto (\sigma_2^2)^{-1}$
0.2	$NIG(3.04, 9.4, 4.2, 389.34)$	$NIG(8.43, 8.8, 3.9, 287.02)$
0.5	$NIG(3.04, 23.5, 11.3, 973.36)$	$NIG(8.43, 22.0, 10.5, 717.55)$
0.8	$NIG(3.04, 37.6, 18.3, 1557.38)$	$NIG(8.43, 35.2, 17.1, 1148.08)$
1.0	$NIG(3.04, 47.0, 23.0, 1946.72)$	$NIG(8.43, 44.0, 21.5, 1435.10)$

Table 5.5: Posterior distributions of (μ_i, σ_i^2).

α	Placebo $(i = 1)$	Drug A $(i = 2)$
0.0	$NIG(3.75, 53.0, 26.0, 1299.61)$	$NIG(10.20, 54.0, 26.5, 2336.56)$
0.2	$NIG(3.64, 62.4, 30.7, 1692.98)$	$NIG(9.95, 62.8, 30.9, 2647.29)$
0.5	$NIG(3.53, 76.5, 37.8, 2281.18)$	$NIG(9.69, 76.0, 37.5, 3103.08)$
0.8	$NIG(3.46, 90.6, 44.8, 2868.07)$	$NIG(9.50, 89.2, 44.1, 3551.40)$
1.0	$NIG(3.42, 100.0, 49.5, 3258.89)$	$NIG(9.41, 98.0, 48.5, 3847.62)$

holds or not. On the other hand, if we cannot pre-specify them, it may be hard to determine the statistical method for comparing the means based on whether $\theta_e(\Delta) \geq \gamma$ or not because it depend on the choice of α, Δ and γ. In such case, we have to determine the statistical method based only on the present trial data, and we can utilize $\theta_e(\Delta)$'s for several α's to scrutinize the appropriateness of the method. Depending on the values of $\theta_e(\Delta)$'s, we may conduct sensitivity analysis by changing the statistical method for comparing the means.

5.6 Conclusion

We have proposed the Bayesian index of superiority to make a more efficient decision for comparing the variances between two groups than possible with the conventional F-test. This index was expressed by the hypergeometric series and the cumulative distribution functions of well-known distributions. Furthermore, we showed that as the amount of prior information decreases, the Bayesian index of superiority approaches the $(1 - p)$ value of the F-test with $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_1 : \sigma_1^2 > \sigma_2^2$. Moreover, if the prior distribution of (μ_i, σ_i^2) is $f(\mu_i, \sigma_i^2) \propto (\sigma_i^2)^{-1}$ for $i = 1, 2$, then $\Pr(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2) = 1 - p$ holds. This indicates that the Bayesian index with a "non-informative" prior or "zero prior effective sample size" can have the same statistical properties as the F-test; however, with incorporation of suitable historical data, the Bayesian index can

Table 5.6: Marginal posterior distributions of σ_i^2 .

α	Placebo $(i = 1)$	Drug A $(i = 2)$
0.0	$Inv\text{-}Ga(26.0, 1299.61)$	$Inv\text{-}Ga(26.5, 2336.56)$
0.2	$Inv\text{-}Ga(30.7, 1692.98)$	$Inv\text{-}Ga(30.9, 2647.29)$
0.5	$Inv\text{-}Ga(37.8, 2281.18)$	$Inv\text{-}Ga(37.5, 3103.08)$
0.8	$Inv\text{-}Ga(44.8, 2868.07)$	$Inv\text{-}Ga(44.1, 3551.40)$
1.0	$Inv\text{-}Ga(49.5, 3258.89)$	$Inv\text{-}Ga(48.5, 3847.62)$

Table 5.7: Bayesian index of superiority and equivalence.

	Superiority		Equivalence			
α	θ	$1-\theta$	$\theta_e(1.10)$	$\theta_e(1.25)$	$\theta_e(1.50)$	$\theta_e(2.00)$
0.0	0.021	0.979	0.084	0.331	0.810	0.998
0.2	0.043	0.957	0.158	0.509	0.926	1.000
0.5	0.087	0.913	0.281	0.715	0.984	1.000
0.8	0.140	0.860	0.403	0.845	0.997	1.000
1.0	0.179	0.821	0.476	0.899	0.999	1.000

potentially be used to make a more efficient decision. In addition, we proposed the Bayesian index of equivalence $\theta_e(\Delta)$, which was evaluated with a Monte Carlo simulation. The results showed that the percentage satisfying $\theta_e(\Delta) \geq \gamma$ heavily depends on the sample size. Therefore, the appropriate values of Δ and γ must be decided on a case-by-case basis. If we mainly focus on comparing the variances, we can utilize the index of superiority and equivalence based on the objectives of trials. If we want to check the assumption about the variances in some statistical method, we can utilize the index of equivalence. In any case, in order to use these indexes for the confirmatory purpose, it is crucial to pre-specify $\alpha_1, \alpha_2, \Delta$, and γ suitably based on the sufficiently reliable information because θ and whether $\theta_e(\Delta) \geq \gamma$ or not depend on them. Therefore, the important future work is to develop a suitable method for constructing the prior distributions, including selecting suitable historical data, and deciding α_1, α_2 for the conditional power prior.

Chapter 6

Bayesian non-inferiority test for two binomial probabilities as the extension of Fisher′s exact test

6.1 Introduction

When approved treatments exist for some diseases or symptoms, non-inferiority trials are planned to show that the new treatment is not worse than an approved treatment by more than a specified margin. Such trials have gained importance in drug and medical device development in recent years. In order to show the non-inferiority, frequentist methods have often been used with $H_0 : \theta_1 \leq \theta_2 - \Delta$ versus $H_1 : \theta_1 > \theta_2 - \Delta$ for some parameters θ_1, θ_2 (e.g., means of normal distributions, probabilities of binomial distributions, rate parameters of Poisson distributions, etc.) and non-inferiority margin $\Delta > 0$. Recently, on the other hand, Bayesian methods have been studied Gamalo et al. (2011, 2016, 2014); Gamalo-Siebers et al. (2016); Ghosh et al. (2016); Kawasaki and Miyaoka (2013); Kawasaki et al. (2016). While the frequentist methods usually use the historical data only for specifying the non-inferiority margin Δ, Bayesian methods can utilize the historical data to specify the prior distributions, thereby enabling more efficient decision making. Bayesian methods often evaluate the posterior probability of H_1 being true, i.e., $\Pr(\theta_1 > \theta_2 - \Delta \mid X_1, X_2)$ where X_1, X_2 are data of the trial treatment group and the active control group, respectively, in present study. The FDA guidance (US Food and Drug Administration and others (2010)) states that “For Bayesian hypothesis testing, you may use the posterior distribution to calculate the probability that a particular hypothesis is true, given the observed data.” in ”5.2 Hypothesis testing”. Therefore, the above procedure can be interpreted as the Bayesian non-inferiority test. For binomial data $X_1 \sim Bin(n_1, \pi_1)$ and $X_2 \sim Bin(n_2, \pi_2)$, Gamalo et al. (2011) evaluated the posterior probability $\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)$ to demonstrate the non-inferiority. However, their calculation method includes the normal approximation and Monte Carlo approximation which are not accurate. Alternatively, Kawasaki and Miyaoka (2013) calculated the integral analytically and expressed it by using the integral of the Appell hypergeometric series, which is quite complicated to calculate. In addition, the relationship between the method in Gamalo et al. (2011) and the frequentist method is not clear. For the superiority test, Kawasaki et al. (2014), Altham (1969) and Howard (1998) derived a clear relationship between the posterior probability and the p-value of the Fisher’s exact test. In this chapter, we extend the method proposed by Gamalo et al. (2011) to the case where the non-inferiority margin Δ is fixed and the historical data can be utilized. First, we derive the exact representation of the posterior probability of H_1 being true under mild conditions. Then, we propose a framework that is more flexible than the method proposed by Gamalo et al. (2011). In our framework, we can incorporate the historical data flexibly by utilizing the conditional power prior and the prior effective sample size. Further, we show the relationship between the posterior probability and the p-value of Fisher’s exact test. From this relationship, we can handle both superiority and non-inferiority in the same framework. We evaluate the proposed method by using Monte Carlo simulations. Additionally, we apply our methods to two HIV clinical trials. Cuffe (2011) stated that these two trials are similar based on the six criteria suggested by Pocock (1976). We treat one of the trials as the historical trial and the other as the present trial. For these trials, however, there are almost four times as much historical data as present data. Therefore, the method proposed in Gamalo et al. (2011) seems to borrow too much information from the historical trial. On the other hand, our method can control the quantity of information borrowed. Finally, we conduct the sample size calculation utilizing the historical data to reduce the sample size of the new trial.

The remainder of this chapter is organized as follows. In Section 6.2, we introduce the Bayesian non-inferiority test with a fixed margin. Next, we describe the proposed method in Section 6.3. Subsequently, we describe the simulations conducted for evaluating the operating characteristics in

Section 6.4. Then, we present a real data analysis and sample size simulations based on the real historical data in Section 6.5. Finally, in Section 6.6, we present our concluding remarks.

6.2 Bayesian non-inferiority test

6.2.1 General configuration

Let X_{1H}, X_{2H} be the random variables corresponding to the responses of trial treatment and the active control group in the past trial, respectively. For $i = 1, 2$, let $X_{iH} \sim Bin(n_{iH}, \pi_{iH})$, and, for the moment, let the prior distribution of π_{iH} be $Beta(\alpha_{0i}, \beta_{0i})$. Later, we allow for some improper priors. Next, in the present trial, let X_1, X_2 be the random variables corresponding to the responses of trial treatment and the active control group, respectively, and suppose $X_i \sim Bin(n_i, \pi_i)$ for $i = 1, 2$.

6.2.2 Gamalo's fully Bayesian method with fixed margin

Here, we introduce the modified version of Gamalo's fully Bayesian method proposed in Gamalo et al. (2011) to the situation where the non-inferiority margin Δ is pre-specified, and the posterior probability is calculated using the Monte Carlo approximation. This method utilizes the data of the active control group from the past trial to construct the prior distribution of the same group in the present trial. Here, the posterior distribution of π_{2H} is given as

$$\pi_{2H} \mid X_{2H} \sim Beta(\alpha_{02} + X_{2H}, \beta_{02} + n_{2H} - X_{2H}).$$

Then, let the prior distribution of π_2 be $Beta(\alpha_{02} + X_{2H}, \beta_{02} + n_{2H} - X_{2H})$. There are no data of the trial treatment in the past trial. Therefore, let the prior distribution of π_1 be $Beta(\alpha_1, \beta_1)$. Then, the posterior distributions of π_1 and π_2 are given as follows:

$$\pi_1 \mid X_1 \sim Beta(\alpha_1 + X_1, \beta_1 + n_1 - X_1),$$
$$\pi_2 \mid X_2 \sim Beta(\alpha_{02} + X_{2H} + X_2, \beta_{02} + n_{2H} - X_{2H} + n_2 - X_2).$$

For the simulation and modified data analysis section in Gamalo et al. (2011), they used $\alpha_1 = \beta_1 = \alpha_{02} = \beta_{02} = 1$. Then, the posterior probability is calculated using the Monte Carlo approximation as follows:

$$\widehat{P}(\pi_1 \geq \pi_2 - \Delta \mid X_1, X_2) = \frac{1}{M} \sum_{i=1}^{M} I(\pi_{1,i} \geq \pi_{2,i} - \Delta)$$

where $\pi_{1,i}$ and $\pi_{2,i}$ for $i = 1, \ldots, M$ are the independent random samples from the posterior distributions of π_1 and π_2, respectively. Here, define the Bayesian decision criterion that the trial treatment is non-inferior to the active control if

$$\widehat{P}(\pi_1 \geq \pi_2 - \Delta \mid X_1, X_2) \geq p^*,$$

where p^* is the pre-specified value. In the simulation section in Gamalo et al. (2011), p^* is assigned a value of 0.81, such that it provides sufficient control of type I error that is not too conservative.

In this method, (i) the probability calculation is based on the Monte Carlo approximation, (ii) past trial data X_{2H} has the same weight as the present trial X_2, and (iii) p^* is determined in a heuristic way. We investigates these points in the following.

6.3 Proposed method

Similar to Gammalo's fully Bayesian method described earlier and the method proposed in Kawasaki and Miyaoka (2013), we evaluate the posterior probability $\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)$ with pre-specified fixed non-inferiority margin Δ. This can be considered as the Bayesian hypothesis testing described in FDA guidance (US Food and Drug Administration and others (2010)). Here, we claim that the trial treatment is non-inferior to the active control if $\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) > p^*$ with pre-specified p^*. In this section, (i) we give the exact expression for the posterior probability $\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)$ under some mild conditions, and (ii) we propose the Bayesian non-inferiority test that can flexibly incorporate the historical data and can be seen as the extension of Fisher's exact test by constructing the prior distributions and deciding the threshold value p^*.

6.3.1 Exact calculation of the posterior probability

First, we consider the exact calculation of the posterior probability $\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)$. Gamalo et al. (2011) used the Monte Carlo simulation or normal approximation, and Kawasaki and Miyaoka (2013) used the normal approximation or the integral of the Appell hypergeometric function. These methods seem insufficient because Monte Carlo simulation is not deterministic, normal approximation is not accurate, and the integral of the Appell hypergeometric function is quite complicated. Here, we derive the simple exact formula.

Theorem 6.1. Suppose that for $i = 1, 2$, the posterior distribution $\pi_i \sim Beta(a_i, b_i)$ with $a_i, b_i \in \mathbb{N}$, then

$$\begin{aligned}
&\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)\\
=&I_\Delta(a_2, b_2)\\
&+\sum_{s_1=0}^{a_1-1}\sum_{s_2=0}^{b_1-1}\binom{a_1+b_1-1}{s_1+s_2}\binom{a_1+b_1-s_1-s_2-2}{a_1-s_1-1}\cdot(-1)^{a_1-s_1-1}\cdot\Delta^{a_1+b_1-1-s_1-s_2}\\
&\qquad\times\binom{s_1+s_2}{s_1}\frac{B(a_2+s_1, b_2+s_2)}{B(a_2,b_2)}\cdot I_{1-\Delta}(b_2+s_2, a_2+s_1)\\
&+\sum_{s_1=0}^{a_1-1}\binom{a_1+b_1-1}{s_1}\frac{B(a_2+s_1, b_2+a_1+b_1-1-s_1)}{B(a_2,b_2)}\\
&\qquad\times I_{1-\Delta}(b_2+a_1+b_1-1-s_1, a_2+s_1),
\end{aligned}\tag{6.1}$$

where

$$I_x(a,b) = \frac{1}{B(a,b)}\int_0^x t^{a-1}(1-t)^{b-1}dt$$

is a regularized incomplete beta function for $0 \le x \le 1$ and $0 < a, b$.

Here, the condition $a_1, b_1, a_2, b_2 \in \mathbb{N}$ is crucial when applying equation (6.2), which relates the beta distribution and binomial distribution.

Proof. We apply the following formula

$$\sum_{r=0}^{k-1}\binom{n}{r}p^r(1-p)^{n-r} = \frac{1}{B(k, n-k+1)}\int_p^1 x^{k-1}(1-x)^{n-k}dx. \tag{6.2}$$

for $0 < p \leq 1$ and $n, k \in \mathbb{N}$. Here,

$$
\begin{aligned}
&\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)\\
=&\int_0^1 \left(\int_{\pi_2-\Delta}^1 \frac{1}{B(a_1,b_1)} \pi_1^{a_1-1}(1-\pi_1)^{b_1-1} \cdot I(0 \leq \pi_1 \leq 1) d\pi_1 \right) \cdot \frac{1}{B(a_2,b_2)} \pi_2^{a_2-1}(1-\pi_2)^{b_2-1} d\pi_2\\
=&\int_0^\Delta \left(\int_{\pi_2-\Delta}^1 \frac{1}{B(a_1,b_1)} \pi_1^{a_1-1}(1-\pi_1)^{b_1-1} \cdot I(0 \leq \pi_1 \leq 1) d\pi_1 \right) \cdot \frac{1}{B(a_2,b_2)} \pi_2^{a_2-1}(1-\pi_2)^{b_2-1} d\pi_2\\
&+ \int_\Delta^1 \left(\int_{\pi_2-\Delta}^1 \frac{1}{B(a_1,b_1)} \pi_1^{a_1-1}(1-\pi_1)^{b_1-1} \cdot I(0 \leq \pi_1 \leq 1) d\pi_1 \right) \cdot \frac{1}{B(a_2,b_2)} \pi_2^{a_2-1}(1-\pi_2)^{b_2-1} d\pi_2,
\end{aligned}
$$

where $I(\cdot)$ is the indicator function. When $\pi_2 < \Delta$,

$$
\int_{\pi_2-\Delta}^1 \frac{1}{B(a_1,b_1)} \pi_1^{a_1-1}(1-\pi_1)^{b_1-1} \cdot I(0 \leq \pi_1 \leq 1) d\pi_1 = \int_0^1 \frac{1}{B(a_1,b_1)} \pi_1^{a_1-1}(1-\pi_1)^{b_1-1} d\pi_1 = 1.
$$

Therefore,

$$
\begin{aligned}
&\int_0^\Delta \left(\int_{\pi_2-\Delta}^1 \frac{1}{B(a_1,b_1)} \pi_1^{a_1-1}(1-\pi_1)^{b_1-1} \cdot I(0 \leq \pi_1 \leq 1) d\pi_1 \right) \cdot \frac{1}{B(a_2,b_2)} \pi_2^{a_2-1}(1-\pi_2)^{b_2-1} d\pi_2\\
=&\int_0^\Delta \frac{1}{B(a_2,b_2)} \pi_2^{a_2-1}(1-\pi_2)^{b_2-1} d\pi_2\\
=&I_\Delta(a_2,b_2).
\end{aligned}
$$

Next, when $\pi_2 \geq \Delta$,

$$
\begin{aligned}
&\int_{\pi_2-\Delta}^1 \frac{1}{B(a_1,b_1)} \pi_1^{a_1-1}(1-\pi_1)^{b_1-1} \cdot I(0 \leq \pi_1 \leq 1) d\pi_1\\
=&\int_{\pi_2-\Delta}^1 \frac{1}{B(a_1,b_1)} \pi_1^{a_1-1}(1-\pi_1)^{b_1-1} d\pi_1\\
=&\sum_{r_1=0}^{a_1-1} \binom{a_1+b_1-1}{r_1} (\pi_2-\Delta)^{r_1} \{1-(\pi_2-\Delta)\}^{a_1+b_1-1-r_1}\\
=&\sum_{r_1=0}^{a_1-1} \binom{a_1+b_1-1}{r_1} \left\{ \sum_{s_1=0}^{r_1} \binom{r_1}{s_1} \pi_2^{s_1} (-\Delta)^{r_1-s_1} \right\}\\
&\times \left\{ \sum_{s_2=0}^{a_1+b_1-1-r_1} \binom{a_1+b_1-1-r_1}{s_2} (1-\pi_2)^{s_2} \Delta^{a_1+b_1-1-r_1-s_2} \right\}
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{r_1=0}^{a_1-1} \sum_{s_1=0}^{r_1} \sum_{s_2=0}^{a_1+b_1-1-r_1} \binom{a_1+b_1-1}{r_1} \binom{r_1}{s_1} \binom{a_1+b_1-1-r_1}{s_2} \\
&\qquad \times (-1)^{r_1-s_1} \Delta^{a_1+b_1-1-s_1-s_2} \cdot \pi_2^{s_1} (1-\pi_2)^{s_2}.
\end{aligned}
$$

Here,

$$
\begin{aligned}
&\int_{\Delta}^{1} \left(\int_{\pi_2-\Delta}^{1} \frac{1}{B(a_1,b_1)} \pi_1^{a_1-1} (1-\pi_1)^{b_1-1} \cdot I(0 \le \pi_1 \le 1) d\pi_1 \right) \\
&\qquad \times \frac{1}{B(a_2,b_2)} \cdot \pi_2^{a_2-1} (1-\pi_2)^{b_2-1} d\pi_2 \\
&= \sum_{r_1=0}^{a_1-1} \sum_{s_1=0}^{r_1} \sum_{s_2=0}^{a_1+b_1-1-r_1} \binom{a_1+b_1-1}{r_1} \binom{r_1}{s_1} \binom{a_1+b_1-1-r_1}{s_2} (-1)^{r_1-s_1} \Delta^{a_1+b_1-1-s_1-s_2} \\
&\quad \times \frac{1}{B(a_2,b_2)} \int_{\Delta}^{1} \pi_2^{a_2+s_1-1} (1-\pi_2)^{b_2+s_2-1} d\pi_2 \\
&= \sum_{r_1=0}^{a_1-1} \sum_{s_1=0}^{r_1} \sum_{s_2=0}^{a_1+b_1-1-r_1} \binom{a_1+b_1-1}{r_1} \binom{r_1}{s_1} \binom{a_1+b_1-1-r_1}{s_2} (-1)^{r_1-s_1} \Delta^{a_1+b_1-1-s_1-s_2} \\
&\quad \times \frac{B(a_2+s_1, b_2+s_2)}{B(a_2,b_2)} \cdot (1 - I_{\Delta}(a_2+s_1, b_2+s_2)) \\
&= \sum_{r_1=0}^{a_1-1} \sum_{s_1=0}^{r_1} \sum_{s_2=0}^{a_1+b_1-1-r_1} \binom{a_1+b_1-1}{r_1} \binom{r_1}{s_1} \binom{a_1+b_1-1-r_1}{s_2} (-1)^{r_1-s_1} \Delta^{a_1+b_1-1-s_1-s_2} \\
&\quad \times \frac{B(a_2+s_1, b_2+s_2)}{B(a_2,b_2)} \cdot I_{1-\Delta}(b_2+s_2, a_2+s_1)
\end{aligned}
$$

From the above, we get

$$
\begin{aligned}
&\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) \\
=& I_{\Delta}(a_2,b_2) + \sum_{r_1=0}^{a_1-1} \sum_{s_1=0}^{r_1} \sum_{s_2=0}^{a_1+b_1-1-r_1} \binom{a_1+b_1-1}{r_1} \binom{r_1}{s_1} \binom{a_1+b_1-1-r_1}{s_2} (-1)^{r_1-s_1} \Delta^{a_1+b_1-1-s_1-s_2} \\
&\qquad \times \frac{B(a_2+s_1, b_2+s_2)}{B(a_2,b_2)} \cdot I_{1-\Delta}(b_2+s_2, a_2+s_1) \\
=& I_{\Delta}(a_2,b_2) + \sum_{s_1=0}^{a_1-1} \sum_{r_1=s_1}^{a_1-1} \sum_{s_2=0}^{a_1+b_1-1-r_1} \binom{a_1+b_1-1}{r_1} \binom{r_1}{s_1} \binom{a_1+b_1-1-r_1}{s_2} (-1)^{r_1-s_1} \Delta^{a_1+b_1-1-s_1-s_2} \\
&\qquad \times \frac{B(a_2+s_1, b_2+s_2)}{B(a_2,b_2)} \cdot I_{1-\Delta}(b_2+s_2, a_2+s_1)
\end{aligned}
$$

$$
\begin{aligned}
=&I_\Delta(a_2,b_2) + \sum_{s_1=0}^{a_1-1}\sum_{s_2=0}^{b_1-1}\sum_{r_1=s_1}^{a_1-1}\binom{a_1+b_1-1}{r_1}\binom{r_1}{s_1}\binom{a_1+b_1-1-r_1}{s_2}(-1)^{r_1-s_1}\Delta^{a_1+b_1-1-s_1-s_2}\\
&\qquad\qquad \times \frac{B(a_2+s_1,b_2+s_2)}{B(a_2,b_2)}\cdot I_{1-\Delta}(b_2+s_2,a_2+s_1)\\
&+\sum_{s_1=0}^{a_1-1}\sum_{s_2=b_1}^{a_1+b_1-1-s_1}\sum_{r_1=s_1}^{a_1+b_1-1-s_2}\binom{a_1+b_1-1}{r_1}\binom{r_1}{s_1}\binom{a_1+b_1-1-r_1}{s_2}(-1)^{r_1-s_1}\Delta^{a_1+b_1-1-s_1-s_2}\\
&\qquad\qquad \times \frac{B(a_2+s_1,b_2+s_2)}{B(a_2,b_2)}\cdot I_{1-\Delta}(b_2+s_2,a_2+s_1)\\
=&I_\Delta(a_2,b_2) + \sum_{s_1=0}^{a_1-1}\sum_{s_2=0}^{b_1-1}\left\{\sum_{r_1=s_1}^{a_1-1}\binom{a_1+b_1-1}{r_1}\binom{r_1}{s_1}\binom{a_1+b_1-1-r_1}{s_2}(-1)^{r_1-s_1}\Delta^{a_1+b_1-1-s_1-s_2}\right\}\\
&\qquad\qquad \times \frac{B(a_2+s_1,b_2+s_2)}{B(a_2,b_2)}\cdot I_{1-\Delta}(b_2+s_2,a_2+s_1)\\
&+\sum_{s_1=0}^{a_1-1}\binom{a_1+b_1-1}{s_1}\frac{B(a_2+s_1,b_2+a_1+b_1-1-s_1)}{B(a_2,b_2)}\cdot I_{1-\Delta}(b_2+a_1+b_1-1-s_1,a_2+s_1)\\
=&I_\Delta(a_2,b_2) + \sum_{s_1=0}^{a_1-1}\sum_{s_2=0}^{b_1-1}\binom{a_1+b_1-1}{s_1+s_2}\binom{a_1+b_1-s_1-s_2-2}{a_1-s_1-1}\cdot(-1)^{a_1-s_1-1}\cdot\Delta^{a_1+b_1-1-s_1-s_2}\\
&\qquad\qquad \times \binom{s_1+s_2}{s_1}\frac{B(a_2+s_1,b_2+s_2)}{B(a_2,b_2)}\cdot I_{1-\Delta}(b_2+s_2,a_2+s_1)\\
&+\sum_{s_1=0}^{a_1-1}\binom{a_1+b_1-1}{s_1}\frac{B(a_2+s_1,b_2+a_1+b_1-1-s_1)}{B(a_2,b_2)}\cdot I_{1-\Delta}(b_2+a_1+b_1-1-s_1,a_2+s_1)
\end{aligned}
$$

□

From the formula 6.1, we obtain the following relations:

Corollary 6.2. If $a_1, b_1, a_2, b_2 \in \mathbb{N}$, then

$$\lim_{\Delta\to+0}\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) = \Pr(\pi_1 > \pi_2 \mid X_1, X_2). \tag{6.3}$$

Proof. From Altham (1969), the right hand side can be expressed as follows:

$$\Pr(\pi_1 > \pi_2 \mid X_1, X_2) = \sum_{r_1=0}^{a_1-1}\binom{a_1+b_1-1}{r_1}\frac{B(a_2+r_1,b_2+a_1+b_1-1-r_1)}{B(a_2,b_2)}.$$

The rest of the proof follows from the direct calculations from formula (6.1). □

Corollary 6.3.

If $a_1, b_1, a_2, b_2 \in \mathbb{N}$ and $\Delta \in \mathbb{Q}$, then

$$\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) \in \mathbb{Q}. \tag{6.4}$$

Proof. In (6.1), since $a_1, b_1, a_2, b_2 \in \mathbb{N}$, $s_1, s_2 \in \{0\} \cup \mathbb{N}$, $0 \leq s_1 \leq a_1 - 1$ and $0 \leq s_2 \leq b_1 - 1$,

$$\binom{a_1 + b_1 - 1}{s_1 + s_2}, \binom{a_1 + b_1 - s_1 - s_2 - 2}{a_1 - s_1 - 1}, \binom{s_1 + s_2}{s_1}, \binom{a_1 + b_1 - 1}{s_1} \in \mathbb{Q}.$$

Since $B(x, y) = \Gamma(x) \cdot \Gamma(y)/\Gamma(x + y)$ and for $x \in \mathbb{N}$, $\Gamma(x) = (x - 1)! \in \mathbb{N}$, therefore

$$B(a_2 + s_1, b_2 + s_2), B(a_2, b_2), B(a_2 + s_1, b_2 + a_1 + b_1 - 1 - s_1) \in \mathbb{Q}.$$

Finally, from 26.5.4 in Abramowitz and Stegun (1964), for $a, n \in \mathbb{N}$ with $n > a$ and $0 < p \leq 1$,

$$I_p(a, n - a + 1) = \sum_{r=a}^{n} \binom{n}{r} p^r (1 - p)^{n-r}.$$

Then, when $a_2, b_2 \in \mathbb{N}$, $s_1, s_2 \in \{0\} \cup \mathbb{N}$ and $\Delta \in \mathbb{Q}$,

$$\Delta^{a_1+b_1-1-s_1-s_2}, I_\Delta(a_2, b_2), I_{1-\Delta}(b_2 + s_2, a_2 + s_1), I_{1-\Delta}(b_2 + a_1 + b_1 - 1 - s_1, a_2 + s_1) \in \mathbb{Q}.$$

From the above, all the terms in (6.1) are in $\mathbb{Q}$, therefore

$$\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) \in \mathbb{Q}.$$

□

The term ”exact” expression is sometimes controversial. However, our expression is no doubt ”exact” because, from the proof of the relation (6.4), the right hand side of (6.1) is the finite sum of the rational numbers.

Our exact expression is useful for decision making in actual clinical trials because the posterior probabilities are accurate and uniquely determined, and the decision making does not depend on the random seeds. However, it takes a little long time to calculate the exact value shown in section

6.4.1. Therefore, in the following, we also use the Monte Carlo approximation when we conduct the Monte Carlo simulations, which, in nature, depend on the random seeds.

6.3.2 Bayesian non-inferiority test as the extension of Fisher′s exact test

Next, we construct the framework of the Bayesian non-inferiority test which can be seen as the extension of Fisher's exact test. In order to construct suitable prior distributions of π_1 and π_2, we first refer to the following theorem shown, at least in a modified form, in Kawasaki et al. (2014), Altham (1969), and Howard (1998). This theorem investigates the relationship between the probability of Bayesian superiority test and the p-value of Fisher's exact test.

Theorem 6.4 (Kawasaki et al. (2014), Altham (1969), Howard (1998)). Suppose that the priors of π_1 and π_2 are $f(\pi_1) \propto \pi_1^{-1}$ and $f(\pi_2) \propto (1-\pi_2)^{-1}$, respectively, and $X_1 > 0, X_2 < n_2$, then between the Bayesian posterior probability Pr$(\pi_1 > \pi_2 \mid X_1, X_2)$ and the one sided p-value of Fisher's exact test with $H_0 : \pi_1 \leq \pi_2$ versus $H_1 : \pi_1 > \pi_2$, the following relation holds

$$\text{Pr}(\pi_1 > \pi_2 \mid X_1, X_2) = 1 - p.$$

Proof. See Kawasaki et al. (2014), Altham (1969), or Howard (1998). □

Next, we consider the conditional power prior discussed by Ibrahim and Chen Ibrahim and Chen (2000). Let us denote the likelihood of π_{iH} as $L(\pi_{iH} \mid X_{iH})$ and the probability density function of the prior of π_{iH} as $f_{0i}(\pi_{iH})$. Then, the "posterior" distribution of π_{iH} is constructed based on the conditional power prior of π_{iH} as

$$f(\pi_{iH} \mid X_{iH}) \propto L(\pi_{iH} \mid X_{iH})^{a_{0i}} f_{0i}(\pi_{iH}),$$

where a_{0i} is a parameter that weighs the historical data relative to the likelihood of the present trial. Then, by replacing π_{iH} by π_i, we can construct the conditional power prior of π_i as

$$f(\pi_i \mid X_{iH}) \propto L(\pi_i \mid X_{iH})^{a_{0i}} f_{0i}(\pi_i).$$

Similar to Theorem 6.4, we define $f_{01}(\pi_1) \propto \pi_1^{-1}$ and $f_{02}(\pi_2) \propto (1-\pi_2)^{-1}$. When $X_{1H} >$

$0, X_{2H} < n_{2H}$, the conditional power prior of π_1 and π_2 are $Beta(a_{01}X_{1H}, a_{01}(n_{1H} - X_{1H}) + 1)$ and $Beta(a_{02}X_{2H}+1, a_{02}(n_{2H}-X_{2H}))$, respectively. However, we cannot use the exact expression for the posterior probability (6.1) for the general beta prior. Therefore, we modify the priors and let all the parameters of beta distribution be natural numbers. Hence, we let

$$\alpha_1 := [a_{01}X_{1H}], \beta_1 := [a_{01}(n_{1H} - X_{1H})] + 1,$$
$$\alpha_2 := [a_{02}X_{2H}] + 1, \beta_2 := [a_{02}(n_{2H} - X_{2H})],$$

where [ ] is the floor function. Finally, we assume the prior of π_1 to be $Beta(\alpha_1, \beta_1)$ if $\alpha_1 > 0$, otherwise it is proportional to $\pi_1^{-1}(1 - \pi_1)^{\beta_1 - 1}$, and assume the prior of π_2 to be $Beta(\alpha_2, \beta_2)$ if $\beta_2 > 0$, otherwise it is proportional to $\pi_2^{\alpha_2 - 1}(1 - \pi_2)^{-1}$. After the present trial, let $a_i := \alpha_i + X_i, b_i := \beta_i + (n_i - X_i)$ for $i = 1, 2$. If $X_1 > 0, X_2 < n_2$, then the posterior of π_i is $Beta(a_i, b_i)$ for $i = 1, 2$. Hence, the following theorem holds.

Theorem 6.5. (i) If $X_1 > 0, X_2 < n_2$, then between $\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)$ and the one sided p-value of Fisher's exact test with $H_0 : \pi_1 \leq \pi_2$ versus $H_1 : \pi_1 > \pi_2$, the following relation holds

$$\lim_{a_{01}, a_{02}, \Delta \to +0} \Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) = 1 - p.$$

(ii) Suppose $\alpha_1, \beta_1, \alpha_2, \beta_2, m_1, m_2 \in \mathbb{N}$, $0 < k_1 \leq m_1$ and $0 \leq k_2 < m_2 - \beta_2$. Then between $\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)$ given $X_1 = k_1 - \alpha_1, n_1 = m_1 - (\alpha_1 + \beta_1) + 1, X_2 = k_2 - \alpha_2 + 1, n_2 = m_2 - (\alpha_2 + \beta_2) + 1$, and the one-sided p-value of Fisher's exact test with $X_1 = k_1, X_2 = k_2, n_1 = m_1, n_2 = m_2$, the following relation holds

$$\lim_{\Delta \to +0} \Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) = 1 - p.$$

Proof. (i) When $X_1 > 0, X_2 < n_2$,

$$a_1 = \alpha_1 + X_1 = [a_{01}X_{1H}] + X_1 \in \mathbb{N},$$
$$b_1 = \beta_1 + (n_1 - X_1) = [a_{01}(n_{1H} - X_{1H})] + 1 + (n_1 - X_1) \in \mathbb{N},$$

$$a_2 = \alpha_2 + X_2 = [a_{02}X_{2H}] + 1 + X_2 \in \mathbb{N},$$
$$b_2 = \beta_2 + (n_2 - X_2) = [a_{02}(n_{2H} - X_{2H})] + (n_2 - X_2) \in \mathbb{N}.$$

Then,

$$\lim_{a_{01},a_{02}\to+0}(a_1, b_1, a_2, b_2) = (X_1, (n_1 - X_1) + 1, X_2 + 1, n_2 - X_2).$$

The right hand side corresponds to the set of parameters of the posterior beta distributions for the priors $f(\pi_1) \propto \pi_1^{-1}$ and $f(\pi_2) \propto (1-\pi_2)^{-1}$. Then, the rest of the proof follows from equation (6.3) and Theorem 6.4.

(ii) When $0 < k_1 \leq m_1$ and $0 \leq k_2 < m_2$,

$$a_1 = \alpha_1 + X_1 = \alpha_1 + (k_1 - \alpha_1) = k_1 \in \mathbb{N},$$
$$b_1 = \beta_1 + (n_1 - X_1) = \beta_1 + \{m_1 - (\alpha_1 + \beta_1) + 1\} - (k_1 - \alpha_1) = m_1 - k_1 + 1 \in \mathbb{N},$$
$$a_2 = \alpha_2 + X_2 = \alpha_2 + (k_2 - \alpha_2 + 1) = k_2 + 1 \in \mathbb{N},$$
$$b_2 = \beta_2 + (n_2 - X_2) = \beta_2 + \{m_2 - (\alpha_2 + \beta_2) + 1\} - (k_2 - \alpha_2 + 1) = m_2 - k_2 \in \mathbb{N}.$$

The rest of the proof is almost the same as (i). □

From Theorem 6.5, we can consider the Bayesian non-inferiority test based on Pr$(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)$ as the extension of Fisher's exact test. Therefore, in the following, we claim that the non-inferiority of the trial treatment holds when Pr$(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) > 0.975$, i.e., $p^* = 0.975$. This corresponds to the one sided Fisher's exact test at significance level 2.5$\%$ when $a_{01}, a_{02}, \Delta \to +0$. Therefore, we can switch between superiority and non-inferiority in the same framework by $\Delta \to +0$. Further, if we choose $a_{02} = 1$, our method reduces to the Gamalo's fully Bayesian method with fixed margin. However, it may borrow too much information from the historical data by $a_{02} = 1$ when the historical trial is larger than the present trial or when the historical trial is not significantly similar to the present trial. For example, see the real data analysis in Section 6.5.1.

6.4 Simulations

In this section, we describe two types of simulations that were conducted. First, we evaluate the accuracy and the computation time of the Monte Carlo approximation. Second, we evaluate the type I error rate and the power of the proposed Bayesian non-inferiority test. In this section, we suppose that historical data exist only for the active control group like that in Gamalo et al. (2011). In the following, let N be the number of iterations, M be the number of the Monte Carlo sampling points, a_{02} be the parameter of the conditional power prior, (n_1, n_2, n_{2H}) be sample sizes, and (p_1, p_2, p_{2H}) be the probabilities of the binomial distributions. Here, let π_1, π_2 be the random variables from the Bayesian viewpoint, and p_1, p_2 be the fixed parameters for generating the simulation data and expressing the hypotheses.

6.4.1 The accuracy of the Monte Carlo approximation and computation time

First, we evaluate the accuracy of the Monte Carlo approximation using the following procedures:

1. Specify N, M, and Δ.

2. Generate $\pi_{1,i} \sim Beta(75, 25), \pi_{2,i} \sim Beta(75, 25)$ for $i = 1, \dots, M$, independently.

3. Calculate the estimated probability

$$\widehat{P}(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) = \frac{1}{M} \sum_{i=1}^{M} I(\pi_{1,i} > \pi_{2,i} - \Delta)$$

   where $I(\cdot)$ is the indicator function.

4. Go back to step 2. and repeat the simulation for N times.

Table 6.1 shows the summary statistics of the Monte Carlo approximations for $N = 10,000$, $M = 10^3, 10^4, 10^5, 10^6, 10^7$, and $\Delta = 0.1$. Here, from the exact expression (6.1), $\Pr(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) = 0.94962$.

Table 6.1: Summary of the result of the Monte Carlo approximations.

Number of sampling	N	Mean	SD	Min	Q1	Median	Q3	Max
10^3	$10,000$	0.94951	0.00692	0.92000	0.94500	0.95000	0.95400	0.97200
10^4	$10,000$	0.94961	0.00219	0.94090	0.94810	0.94970	0.95110	0.95860
10^5	$10,000$	0.94963	0.00070	0.94710	0.94916	0.94961	0.95010	0.95218
10^6	$10,000$	0.94962	0.00022	0.94883	0.94947	0.94962	0.94977	0.95050
10^7	$10,000$	0.94962	0.00007	0.94936	0.94957	0.94962	0.94967	0.94989

Next, we consider the computation time to calculate the posterior probability. We performed the simulations using R 3.2.3 on Intel(R) Core(TM) i5-6200U CPU @ 2.30 GHz; RAM 4.00 GB and 64 bit operating system machine. In order to calculate the exact expression, we utilized R package Rmpfr. Table 6.2 shows the computation time required to calculate the posterior probability by the Monte Carlo approximation for $N = 1$ and $M = 10^3, 10^4, 10^5, 10^6, 10^7$, and by the exact expression (6.1).

Table 6.2: Summary of the calculation time.

	Number of sampling of the Monte Carlo approximations					Exact
	10^3	10^4	10^5	10^6	10^7	
Time to calculate one probability (sec)	0.017	0.020	0.081	1.261	7.527	17.698

Taking the accuracy and computation time into consideration, we choose $M = 10^6$ for the following Monte Carlo simulations.

6.4.2 Operating characteristics of the Bayesian non-inferiority test (setting)

Here, we evaluate the type I error rate and the power of our proposed method. As stated above, we suppose that historical data exist only for the active control group. The following procedures are used.

The simulation procedure is as follows:

1. Specify N, M, Δ, (n_{2H}, n_1, n_2), a_{02}, and (p_{2H}, p_1, p_2). Set COUNT = 0.

2. Generate the data of the historical trial $X_{2H} \sim Bin(n_{2H}, p_{2H})$.

3. For $\alpha_2 := [a_{02}X_{2H}] + 1$ and $\beta_2 := [a_{02}(n_{2H} - X_{2H})]$, let the prior distribution of π_2 be $Beta(\alpha_2, \beta_2)$ if $\beta_2 > 0$, else let it be proportional to $\pi_2^{\alpha_2 - 1}(1 - \pi_2)^{-1}$. Let the prior of π_1 be proportional to π_1^{-1}.

4. Generate the data of the present trial $X_1 \sim Bin(n_1, p_1)$, and $X_2 \sim Bin(n_2, p_2)$. In the following suppose $X_1 > 0$ and $X_2 < n_2$.

5. Generate the independent samples of $\pi_{1,i}, \pi_{2,i}$ for $i = 1, \ldots, M$ from the posterior distributions of π_1 and π_2:

$$\pi_{1,i} \sim Beta(a_1, b_1), \pi_{2,i} \sim Beta(a_2, b_2),$$

   where $a_1 := X_1, b_1 := (n_1 - X_1) + 1, a_2 := \alpha_2 + X_2, b_2 := \beta_2 + (n_2 - X_2)$.

6. Calculate the estimated posterior probability

$$\widehat{P}(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) = \frac{1}{M} \sum_{i=1}^{M} I(\pi_{1,i} > \pi_{2,i} - \Delta).$$

7. If $\widehat{P}(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) > 0.975$, then increase the COUNT by 1; otherwise, by 0.

8. Go back to step 2. and repeat the simulation for N times.

9. Calculate type I error rate or power by COUNTS$/N$.

In each iteration above, we also consider the frequentist non-inferiority test whose null and alternative hypotheses are $H_0 : p_1 \leq p_2 - \Delta$ and $H_1 : p_1 > p_2 - \Delta$, respectively. We calculate the proportion by which the lower limits of the two sided 95$\%$ Wald confidence intervals of $p_1 - p_2$ exceed $-\Delta$.

The parameters are as follows:

- $N = 10,000, M = 10^6, \Delta = 0.10.$

- $a_{02} = 0, 0.1, 0.25, 0.5, 0.75, 1$.
- For the type I error rate,
  - Sample sizes:
    * Balanced $(n_1 = n_2)$ case, $(n_{2H}, n_1, n_2) = (100, 100, 100)$, $(200, 200, 200)$, $(300, 200, 200)$, $(400, 200, 200)$, $(300, 300, 300)$, $(400, 400, 400)$.
    * Unbalanced $(n_1 > n_2)$ case, $(n_{2H}, n_1, n_2) = (100, 100, 50)$, $(200, 200, 100)$, $(300, 200, 100)$, $(400, 200, 100)$, $(300, 300, 100)$, $(300, 300, 150)$, $(400, 400, 100)$, $(400, 400, 200)$.
  - Probabilities:
    * $p_1 = 0.60, p_2 = 0.70$ and $p_{2H} = 0.65, 0.70, 0.75$.
- For the power,
  - Sample sizes:
    * Balanced $(n_1 = n_2)$ case, $(n_{2H}, n_1, n_2) = (100, 100, 100)$, $(200, 200, 200)$, $(400, 200, 200)$, $(300, 300, 300)$, $(400, 400, 400)$.
    * Unbalanced $(n_1 > n_2)$ case, $(n_{2H}, n_1, n_2) = (100, 100, 50)$, $(200, 200, 100)$, $(400, 200, 100)$, $(300, 300, 150)$, $(400, 400, 100)$.
  - Probabilities:
    * (Scenario 1) $p_{2H} = p_2 = 0.70$, and p_1 are from 0.62 to 0.80 by increments of 0.02.
    * (Scenario 2) $p_{2H} = 0.75 > p_2 = 0.70$, and p_1 are from 0.62 to 0.80 by increments of 0.02.
    * (Scenario 3) $p_{2H} = 0.65 < p_2 = 0.70$, and p_1 are from 0.62 to 0.80 by increments of 0.02.

6.4.3 Operating characteristics of the Bayesian non-inferiority test (result)

The results of the type I error rate simulations are shown in Table 6.3. First, for the frequentist and Bayesian methods with $a_{02} = 0$, type I error rates are almost always less than 0.025. Next, consider the Bayesian methods with $a_{02} > 0$. When the distribution of the historical data is the same as that of the present one $(p_{2H} = p_2 = 0.70)$, the type I error rates are almost less than 0.025. Further, for many situations especially $0 < a_{02} < 1$, they are more conservative than the frequentist method. For fixed a_{02}, sample size and imbalance do not significantly affect the type I error rate. When the response rate of the historical data is less than that of the present one $(p_{2H} = 0.65 < p_2 = 0.70)$, the type I error rates of the Bayesian method (we call the Bayesian type I error rates, in the following) often exceed 0.025. Inflations tend to be large when a_{02} or n_{2H} are large. Further, by comparing, for example, $(n_{2H}, n_1, n_2) = (300, 300, 100), (300, 300, 150), (300, 300, 300)$, inflation tends to be greater when n_2 is small, i.e., when the weight of the historical trial is large. Conversely, when the response rate of the historical data is more than that of the present one $(p_{2H} = 0.75 > p_2 = 0.70)$, the Bayesian type I error rates tend to be too conservative. In particular, they tend to be more conservative when the sample sizes of the historical trial n_{2H} are large.

Table 6.3: Type I error rate of the Bayesian and frequentist method at different levels of p_{2H} and sample sizes $(p_1 = 0.60, p_2 = 0.70, \Delta = 0.10)$.

p_{2H}	n_{2H}	n_1	n_2	Frequentist	Bayesian method					
					$a_{02} = 0$	0.1	0.25	0.50	0.75	1.0
0.70	100	100	100	0.0246	0.0176	0.0156	0.0152	0.0156	0.0159	0.0180
	200	200	200	0.0250	0.0206	0.0173	0.0158	0.0172	0.0186	0.0204
	300	200	200	0.0232	0.0189	0.0156	0.0147	0.0161	0.0178	0.0208
	400	200	200	0.0257	0.0201	0.0169	0.0159	0.0168	0.0187	0.0207
	300	300	300	0.0235	0.0201	0.0170	0.0161	0.0164	0.0169	0.0185
	400	400	400	0.0232	0.0197	0.0184	0.0180	0.0186	0.0203	0.0231
	100	100	50	0.0230	0.0165	0.0118	0.0116	0.0143	0.0165	0.0198
	200	200	100	0.0232	0.0189	0.0146	0.0126	0.0134	0.0153	0.0187
	300	200	100	0.0254	0.0204	0.0139	0.0120	0.0140	0.0171	0.0214
	400	200	100	0.0204	0.0168	0.0105	0.0102	0.0121	0.0152	0.0181
	300	300	100	0.0214	0.0181	0.0123	0.0116	0.0145	0.0180	0.0221
	300	300	150	0.0211	0.0178	0.0146	0.0123	0.0153	0.0186	0.0205
	400	400	100	0.0242	0.0214	0.0118	0.0118	0.0135	0.0172	0.0226
	400	400	200	0.0299	0.0271	0.0188	0.0172	0.0175	0.0198	0.0241
0.65	100	100	100	0.0266	0.0178	0.0195	0.0230	0.0310	0.0399	0.0516
	200	200	200	0.0235	0.0185	0.0204	0.0268	0.0409	0.0573	0.0742
	300	200	200	0.0304	0.0242	0.0282	0.0383	0.0574	0.0757	0.0978
	400	200	200	0.0255	0.0210	0.0275	0.0412	0.0656	0.0915	0.1136
	300	300	300	0.0247	0.0214	0.0242	0.0318	0.0503	0.0748	0.0993
	400	400	400	0.0244	0.0210	0.0257	0.0356	0.0577	0.0846	0.1152
	100	100	50	0.0283	0.0213	0.0197	0.0229	0.0349	0.0467	0.0630
	200	200	100	0.0247	0.0209	0.0223	0.0334	0.0536	0.0758	0.1012
	300	200	100	0.0244	0.0211	0.0239	0.0372	0.0680	0.0993	0.1260
	400	200	100	0.0224	0.0186	0.0245	0.0413	0.0761	0.1061	0.1345
	300	300	100	0.0202	0.0179	0.0219	0.0363	0.0766	0.1175	0.1545
	300	300	150	0.0245	0.0212	0.0255	0.0377	0.0676	0.1001	0.1374
	400	400	100	0.0210	0.0179	0.0233	0.0453	0.1030	0.1573	0.2089
	400	400	200	0.0241	0.0212	0.0267	0.0430	0.0820	0.1266	0.1674
0.75	100	100	100	0.0255	0.0182	0.0135	0.0096	0.0082	0.0075	0.0073
	200	200	200	0.0268	0.0214	0.0143	0.0099	0.0067	0.0053	0.0045
	300	200	200	0.0227	0.0178	0.0101	0.0060	0.0037	0.0028	0.0023
	400	200	200	0.0272	0.0210	0.0104	0.0056	0.0039	0.0024	0.0018
	300	300	300	0.0257	0.0217	0.0149	0.0096	0.0053	0.0034	0.0029
	400	400	400	0.0258	0.0216	0.0138	0.0088	0.0040	0.0025	0.0023
	100	100	50	0.0236	0.0174	0.0108	0.0064	0.0049	0.0049	0.0050
	200	200	100	0.0231	0.0188	0.0083	0.0041	0.0026	0.0022	0.0025
	300	200	100	0.0236	0.0189	0.0065	0.0022	0.0019	0.0016	0.0018
	400	200	100	0.0236	0.0191	0.0052	0.0022	0.0011	0.0008	0.0008
	300	300	100	0.0208	0.0177	0.0055	0.0016	0.0010	0.0009	0.0013
	300	300	150	0.0232	0.0214	0.0108	0.0052	0.0029	0.0017	0.0016
	400	400	100	0.0232	0.0206	0.0046	0.0009	0.0005	0.0006	0.0005
	400	400	200	0.0245	0.0215	0.0095	0.0034	0.0010	0.0006	0.0005

Tables 6.4 and 6.5 show the results of the power simulations with $p_2 = p_{2H} = 0.70$ for the

balanced $(n_1 = n_2)$ and unbalanced $(n_1 > n_2)$ case, respectively. When p_1 and a_{02} are small, the powers of the our proposed method (also referred to as Bayesian powers) are less than the powers of the frequentist method (also referred to as frequentist powers). On the other hand, when they are large, the converse holds. For the unbalanced case, when compared to the balanced case, the Bayesian powers tend to exceed the frequentist ones even with a smaller value of a_{02}.

Table 6.4: Power of the Bayesian and frequentist method at different levels of p_1 and sample sizes (scenario 1: $p_{2H} = p_2 = 0.70, \Delta = 0.10$, balanced).

p_{2H}	p_1	n_{2H}	n_1	n_2	Frequentist	Bayesian method					
						$a_{02} = 0$	0.1	0.25	0.50	0.75	1.0
	0.62	100	100	100	0.0469	0.0356	0.0332	0.0330	0.0340	0.0357	0.0400
	0.64	100	100	100	0.0866	0.0673	0.0641	0.0625	0.0671	0.0724	0.0804
	0.66	100	100	100	0.1469	0.1149	0.1124	0.1136	0.1222	0.1280	0.1419
	0.68	100	100	100	0.2420	0.1985	0.1953	0.2021	0.2144	0.2240	0.2433
	0.70	100	100	100	0.3494	0.2965	0.2968	0.3079	0.3279	0.3472	0.3674
	0.72	100	100	100	0.4777	0.4139	0.4194	0.4406	0.4722	0.4954	0.5172
	0.74	100	100	100	0.6016	0.5366	0.5479	0.5764	0.6098	0.6368	0.6623
	0.76	100	100	100	0.7198	0.6621	0.6768	0.7057	0.7392	0.7642	0.7799
	0.78	100	100	100	0.8255	0.7825	0.7983	0.8251	0.8509	0.8714	0.8827
	0.80	100	100	100	0.9093	0.8738	0.8880	0.9103	0.9285	0.9386	0.9438
	0.62	200	200	200	0.0625	0.0505	0.0475	0.0452	0.0455	0.0481	0.0531
	0.64	200	200	200	0.1322	0.1118	0.1083	0.1116	0.1192	0.1274	0.1382
	0.66	200	200	200	0.2533	0.2233	0.2250	0.2349	0.2560	0.2685	0.2874
	0.68	200	200	200	0.4098	0.3714	0.3778	0.3944	0.4239	0.4432	0.4664
	0.70	200	200	200	0.5875	0.5525	0.5625	0.5837	0.6232	0.6454	0.6649
	0.72	200	200	200	0.7519	0.7164	0.7340	0.7648	0.7975	0.8164	0.8297
	0.74	200	200	200	0.8785	0.8526	0.8672	0.8929	0.9146	0.9249	0.9333
	0.76	200	200	200	0.9537	0.9429	0.9529	0.9647	0.9749	0.9782	0.9808
	0.78	200	200	200	0.9847	0.9807	0.9845	0.9892	0.9930	0.9943	0.9952
	0.80	200	200	200	0.9955	0.9940	0.9963	0.9978	0.9984	0.9987	0.9990
	0.62	400	200	200	0.0605	0.0501	0.0471	0.0477	0.0527	0.0581	0.0631
	0.64	400	200	200	0.1340	0.1162	0.1136	0.1178	0.1301	0.1395	0.1519
	0.66	400	200	200	0.2475	0.2183	0.2220	0.2388	0.2678	0.2909	0.3095
	0.68	400	200	200	0.4127	0.3768	0.3973	0.4261	0.4692	0.4951	0.5162
0.70	0.70	400	200	200	0.5841	0.5462	0.5762	0.6212	0.6694	0.6929	0.7086
	0.72	400	200	200	0.7541	0.7246	0.7574	0.7984	0.8379	0.8569	0.8699
	0.74	400	200	200	0.8765	0.8559	0.8846	0.9159	0.9382	0.9487	0.9545
	0.76	400	200	200	0.9522	0.9421	0.9595	0.9739	0.9841	0.9865	0.9877
	0.78	400	200	200	0.9825	0.9783	0.9882	0.9934	0.9963	0.9977	0.9982
	0.80	400	200	200	0.9967	0.9954	0.9983	0.9994	0.9996	0.9998	0.9998
	0.62	300	300	300	0.0753	0.0653	0.0617	0.0620	0.0639	0.0673	0.0734
	0.64	300	300	300	0.1800	0.1588	0.1576	0.1663	0.1751	0.1875	0.2004
	0.66	300	300	300	0.3487	0.3189	0.3266	0.3449	0.3719	0.3923	0.4146
	0.68	300	300	300	0.5653	0.5308	0.5449	0.5741	0.6061	0.6310	0.6509
	0.70	300	300	300	0.7684	0.7411	0.7619	0.7884	0.8150	0.8348	0.8484
	0.72	300	300	300	0.9074	0.8909	0.9064	0.9228	0.9399	0.9482	0.9545
	0.74	300	300	300	0.9704	0.9638	0.9704	0.9783	0.9841	0.9865	0.9887
	0.76	300	300	300	0.9913	0.9892	0.9922	0.9961	0.9979	0.9984	0.9986
	0.78	300	300	300	0.9982	0.9980	0.9985	0.9991	0.9995	0.9999	0.9999
	0.80	300	300	300	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	0.62	400	400	400	0.0863	0.0736	0.0714	0.0736	0.0780	0.0840	0.0907
	0.64	400	400	400	0.2297	0.2100	0.2109	0.2157	0.2307	0.2444	0.2602
	0.66	400	400	400	0.4422	0.4149	0.4225	0.4475	0.4769	0.4987	0.5181
	0.68	400	400	400	0.6884	0.6620	0.6788	0.7075	0.7382	0.7581	0.7747
	0.70	400	400	400	0.8710	0.8550	0.8715	0.8921	0.9135	0.9257	0.9338
	0.72	400	400	400	0.9623	0.9561	0.9646	0.9735	0.9815	0.9844	0.9862
	0.74	400	400	400	0.9943	0.9930	0.9951	0.9971	0.9984	0.9989	0.9992
	0.76	400	400	400	0.9992	0.9990	0.9994	0.9998	0.9999	0.9999	0.9999
	0.78	400	400	400	0.9999	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000
	0.80	400	400	400	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 6.5: Power of the Bayesian and frequentist method at different levels of p_1 and sample sizes (scenario 1: $p_{2H} = p_2 = 0.70, \Delta = 0.10$, unbalanced).

p_{2H}	p_1	n_{2H}	n_1	n_2	Frequentist	Bayesian method					
						$a_{02} = 0$	0.1	0.25	0.50	0.75	1.0
	0.62	100	100	50	0.0375	0.0290	0.0250	0.0241	0.0278	0.0318	0.0392
	0.64	100	100	50	0.0687	0.0550	0.0505	0.0498	0.0580	0.0639	0.0777
	0.66	100	100	50	0.1044	0.0884	0.0835	0.0876	0.1031	0.1175	0.1355
	0.68	100	100	50	0.1679	0.1436	0.1418	0.1499	0.1845	0.2072	0.2362
	0.70	100	100	50	0.2298	0.2021	0.2073	0.2286	0.2723	0.3039	0.3385
	0.72	100	100	50	0.3293	0.2937	0.3095	0.3451	0.4018	0.4418	0.4760
	0.74	100	100	50	0.4253	0.3908	0.4130	0.4650	0.5310	0.5728	0.6106
	0.76	100	100	50	0.5526	0.5125	0.5490	0.6055	0.6713	0.7105	0.7428
	0.78	100	100	50	0.6637	0.6278	0.6698	0.7295	0.7964	0.8314	0.8547
	0.80	100	100	50	0.7639	0.7331	0.7776	0.8330	0.8868	0.9098	0.9233
	0.62	200	200	100	0.0514	0.0437	0.0365	0.0347	0.0378	0.0433	0.0541
	0.64	200	200	100	0.0999	0.0871	0.0817	0.0877	0.1009	0.1140	0.1304
	0.66	200	200	100	0.1708	0.1533	0.1516	0.1696	0.1998	0.2266	0.2530
	0.68	200	200	100	0.2904	0.2655	0.2734	0.3088	0.3604	0.3969	0.4355
	0.70	200	200	100	0.4220	0.3948	0.4186	0.4809	0.5520	0.5874	0.6209
	0.72	200	200	100	0.5808	0.5525	0.5910	0.6646	0.7357	0.7700	0.7982
	0.74	200	200	100	0.7151	0.6947	0.7385	0.8073	0.8616	0.8880	0.9041
	0.76	200	200	100	0.8434	0.8292	0.8699	0.9205	0.9507	0.9623	0.9683
	0.78	200	200	100	0.9219	0.9129	0.9452	0.9717	0.9867	0.9905	0.9925
	0.80	200	200	100	0.9676	0.9628	0.9810	0.9913	0.9959	0.9974	0.9983
	0.62	400	200	100	0.0457	0.0378	0.0309	0.0320	0.0402	0.0483	0.0552
	0.64	400	200	100	0.0948	0.0796	0.0735	0.0837	0.1094	0.1256	0.1421
	0.66	400	200	100	0.1709	0.1514	0.1592	0.1948	0.2456	0.2793	0.3026
	0.68	400	200	100	0.2819	0.2571	0.2898	0.3514	0.4249	0.4627	0.4905
	0.70	400	200	100	0.4294	0.4017	0.4673	0.5472	0.6276	0.6675	0.6926
0.70	0.72	400	200	100	0.5705	0.5446	0.6335	0.7325	0.8031	0.8316	0.8501
	0.74	400	200	100	0.7255	0.7026	0.8005	0.8750	0.9214	0.9376	0.9473
	0.76	400	200	100	0.8441	0.8294	0.9053	0.9523	0.9753	0.9822	0.9851
	0.78	400	200	100	0.9223	0.9140	0.9676	0.9873	0.9956	0.9977	0.9982
	0.80	400	200	100	0.9656	0.9610	0.9886	0.9973	0.9990	0.9993	0.9996
	0.62	300	300	150	0.0595	0.0520	0.0476	0.0453	0.0533	0.0623	0.0723
	0.64	300	300	150	0.1287	0.1168	0.1126	0.1167	0.1346	0.1611	0.1835
	0.66	300	300	150	0.2497	0.2323	0.2383	0.2623	0.3117	0.3498	0.3832
	0.68	300	300	150	0.4054	0.3854	0.4157	0.4664	0.5361	0.5783	0.6093
	0.70	300	300	150	0.5812	0.5636	0.6026	0.6706	0.7420	0.7803	0.8047
	0.72	300	300	150	0.7532	0.7397	0.7849	0.8453	0.8963	0.9171	0.9290
	0.74	300	300	150	0.8927	0.8825	0.9173	0.9509	0.9708	0.9788	0.9822
	0.76	300	300	150	0.9567	0.9520	0.9724	0.9882	0.9952	0.9972	0.9981
	0.78	300	300	150	0.9853	0.9837	0.9918	0.9979	0.9992	0.9994	0.9998
	0.80	300	300	150	0.9966	0.9964	0.9987	0.9994	0.9999	1.0000	1.0000
	0.62	400	400	100	0.0525	0.0483	0.0371	0.0389	0.0546	0.0692	0.0822
	0.64	400	400	100	0.1075	0.1005	0.0965	0.1173	0.1580	0.1931	0.2254
	0.66	400	400	100	0.2060	0.1989	0.2151	0.2692	0.3553	0.4129	0.4549
	0.68	400	400	100	0.3352	0.3284	0.3856	0.4887	0.6099	0.6681	0.7077
	0.70	400	400	100	0.4980	0.4934	0.5947	0.7213	0.8192	0.8646	0.8878
	0.72	400	400	100	0.6576	0.6552	0.7738	0.8845	0.9471	0.9650	0.9714
	0.74	400	400	100	0.8037	0.8043	0.9100	0.9688	0.9892	0.9940	0.9956
	0.76	400	400	100	0.9037	0.9056	0.9730	0.9939	0.9993	0.9998	0.9998
	0.78	400	400	100	0.9657	0.9663	0.9936	0.9997	1.0000	1.0000	1.0000
	0.80	400	400	100	0.9878	0.9884	0.9994	1.0000	1.0000	1.0000	1.0000

Tables 6.6 and 6.7 show the results of the power simulations with $p_{2H} = 0.75 > p_2 = 0.70$ for the balanced $(n_1 = n_2)$ and unbalanced $(n_1 > n_2)$ case, respectively. In these cases, the frequentist powers are almost always more than the Bayesian powers for all value of $0.62 \leq p_1 \leq 0.80$. For the unbalanced case, the Bayesian powers sometimes exceed the frequentist powers where $0.74 \leq p_1$.

Table 6.6: Power of the Bayesian and frequentist method at different levels of p_1 and sample sizes (scenario 2: $p_{2H} = 0.75 > p_2 = 0.70, \Delta = 0.10$, balanced).

p_{2H}	p_1	n_{2H}	n_1	n_2	Frequentist	Bayesian method					
						$a_{02} = 0$	0.1	0.25	0.50	0.75	1.0
	0.62	100	100	100	0.0514	0.0376	0.0277	0.0239	0.0173	0.0144	0.0140
	0.64	100	100	100	0.0933	0.0741	0.0605	0.0504	0.0417	0.0376	0.0363
	0.66	100	100	100	0.1484	0.1153	0.1007	0.0855	0.0757	0.0697	0.0697
	0.68	100	100	100	0.2346	0.1909	0.1708	0.1572	0.1410	0.1332	0.1328
	0.70	100	100	100	0.3456	0.2965	0.2697	0.2522	0.2354	0.2274	0.2247
	0.72	100	100	100	0.4699	0.4100	0.3885	0.3769	0.3628	0.3525	0.3487
	0.74	100	100	100	0.6024	0.5441	0.5285	0.5163	0.5039	0.4974	0.4927
	0.76	100	100	100	0.7253	0.6700	0.6620	0.6589	0.6513	0.6482	0.6426
	0.78	100	100	100	0.8366	0.7912	0.7864	0.7895	0.7862	0.7831	0.7803
	0.80	100	100	100	0.9107	0.8813	0.8820	0.8906	0.8915	0.8940	0.8916
	0.62	200	200	200	0.0645	0.0525	0.0405	0.0302	0.0208	0.0161	0.0139
	0.64	200	200	200	0.1338	0.1109	0.0936	0.0751	0.0615	0.0521	0.0462
	0.66	200	200	200	0.2479	0.2182	0.1912	0.1659	0.1399	0.1239	0.1145
	0.68	200	200	200	0.3994	0.3672	0.3364	0.3085	0.2763	0.2550	0.2444
	0.70	200	200	200	0.5900	0.5519	0.5272	0.5013	0.4723	0.4466	0.4326
	0.72	200	200	200	0.7597	0.7286	0.7125	0.6995	0.6766	0.6574	0.6465
	0.74	200	200	200	0.8807	0.8596	0.8527	0.8500	0.8411	0.8292	0.8209
	0.76	200	200	200	0.9485	0.9369	0.9353	0.9365	0.9338	0.9275	0.9254
	0.78	200	200	200	0.9856	0.9812	0.9819	0.9832	0.9834	0.9823	0.9809
	0.80	200	200	200	0.9961	0.9948	0.9953	0.9957	0.9957	0.9957	0.9957
	0.62	400	200	200	0.0607	0.0508	0.0309	0.0195	0.0132	0.0098	0.0087
	0.64	400	200	200	0.1329	0.1142	0.0799	0.0556	0.0405	0.0336	0.0308
	0.66	400	200	200	0.2558	0.2232	0.1779	0.1389	0.1094	0.0942	0.0872
	0.68	400	200	200	0.4132	0.3793	0.3225	0.2757	0.2363	0.2128	0.2023
0.75	0.70	400	200	200	0.5941	0.5535	0.5114	0.4692	0.4278	0.4070	0.3920
	0.72	400	200	200	0.7556	0.7246	0.6977	0.6761	0.6453	0.6198	0.6069
	0.74	400	200	200	0.8769	0.8548	0.8423	0.8345	0.8198	0.8054	0.7959
	0.76	400	200	200	0.9500	0.9386	0.9372	0.9365	0.9315	0.9263	0.9206
	0.78	400	200	200	0.9851	0.9804	0.9806	0.9815	0.9799	0.9780	0.9758
	0.80	400	200	200	0.9963	0.9948	0.9958	0.9963	0.9963	0.9961	0.9963
	0.62	300	300	300	0.0772	0.0664	0.0469	0.0329	0.0224	0.0167	0.0148
	0.64	300	300	300	0.1796	0.1583	0.1304	0.1049	0.0793	0.0641	0.0571
	0.66	300	300	300	0.3540	0.3239	0.2869	0.2517	0.2120	0.1816	0.1671
	0.68	300	300	300	0.5643	0.5286	0.4942	0.4582	0.4119	0.3790	0.3614
	0.70	300	300	300	0.7500	0.7240	0.7039	0.6847	0.6487	0.6175	0.5985
	0.72	300	300	300	0.8984	0.8818	0.8724	0.8642	0.8472	0.8307	0.8191
	0.74	300	300	300	0.9674	0.9610	0.9599	0.9570	0.9524	0.9468	0.9422
	0.76	300	300	300	0.9920	0.9890	0.9894	0.9895	0.9895	0.9884	0.9862
	0.78	300	300	300	0.9989	0.9987	0.9988	0.9989	0.9990	0.9990	0.9988
	0.80	300	300	300	1.0000	1.0000	0.9999	0.9999	0.9999	0.9999	0.9999
	0.62	400	400	400	0.0820	0.0724	0.0555	0.0399	0.0255	0.0185	0.0144
	0.64	400	400	400	0.2330	0.2116	0.1706	0.1341	0.0965	0.0766	0.0671
	0.66	400	400	400	0.4499	0.4230	0.3770	0.3286	0.2755	0.2349	0.2142
	0.68	400	400	400	0.6944	0.6678	0.6311	0.5922	0.5416	0.4977	0.4691
	0.70	400	400	400	0.8633	0.8473	0.8324	0.8124	0.7857	0.7557	0.7347
	0.72	400	400	400	0.9647	0.9589	0.9551	0.9487	0.9398	0.9292	0.9186
	0.74	400	400	400	0.9931	0.9915	0.9914	0.9919	0.9902	0.9884	0.9862
	0.76	400	400	400	0.9988	0.9983	0.9984	0.9984	0.9983	0.9983	0.9981
	0.78	400	400	400	0.9999	0.9998	0.9999	0.9999	1.0000	1.0000	1.0000
	0.80	400	400	400	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 6.7: Power of the Bayesian and frequentist method at different levels of p_1 and sample sizes (scenario 2: $p_{2H} = 0.75 > p_2 = 0.70, \Delta = 0.10$, unbalanced)

p_{2H}	p_1	n_{2H}	n_1	n_2	Frequentist	Bayesian method					
						$a_{02} = 0$	0.1	0.25	0.50	0.75	1.0
0.75	0.62	100	100	50	0.0396	0.0306	0.0213	0.0143	0.0113	0.0104	0.0114
	0.64	100	100	50	0.0681	0.0558	0.0391	0.0277	0.0235	0.0229	0.0241
	0.66	100	100	50	0.1013	0.0846	0.0658	0.0545	0.0493	0.0479	0.0520
	0.68	100	100	50	0.1576	0.1344	0.1116	0.0942	0.0895	0.0893	0.0942
	0.70	100	100	50	0.2238	0.1970	0.1684	0.1574	0.1569	0.1595	0.1697
	0.72	100	100	50	0.3146	0.2802	0.2548	0.2503	0.2539	0.2573	0.2713
	0.74	100	100	50	0.4255	0.3871	0.3685	0.3694	0.3769	0.3852	0.4028
	0.76	100	100	50	0.5453	0.5070	0.4962	0.5042	0.5205	0.5326	0.5489
	0.78	100	100	50	0.6625	0.6282	0.6294	0.6475	0.6775	0.6893	0.7057
	0.80	100	100	50	0.7683	0.7389	0.7488	0.7675	0.8012	0.8139	0.8248
	0.62	200	200	100	0.0458	0.0380	0.0241	0.0161	0.0105	0.0091	0.0095
	0.64	200	200	100	0.0993	0.0878	0.0607	0.0422	0.0313	0.0276	0.0282
	0.66	200	200	100	0.1754	0.1552	0.1201	0.0985	0.0856	0.0778	0.0791
	0.68	200	200	100	0.2907	0.2610	0.2184	0.1960	0.1776	0.1693	0.1733
	0.70	200	200	100	0.4235	0.3948	0.3571	0.3432	0.3272	0.3176	0.3207
	0.72	200	200	100	0.5820	0.5566	0.5311	0.5300	0.5238	0.5154	0.5206
	0.74	200	200	100	0.7259	0.7040	0.6946	0.7084	0.7120	0.7090	0.7169
	0.76	200	200	100	0.8367	0.8204	0.8253	0.8437	0.8521	0.8542	0.8590
	0.78	200	200	100	0.9238	0.9143	0.9244	0.9393	0.9491	0.9499	0.9516
	0.80	200	200	100	0.9680	0.9626	0.9707	0.9817	0.9869	0.9873	0.9880
	0.62	400	200	100	0.0477	0.0398	0.0166	0.0091	0.0065	0.0051	0.0058
	0.64	400	200	100	0.1017	0.0876	0.0470	0.0309	0.0235	0.0235	0.0235
	0.66	400	200	100	0.1766	0.1562	0.1032	0.0735	0.0634	0.0609	0.0613
	0.68	400	200	100	0.2840	0.2587	0.2005	0.1720	0.1603	0.1556	0.1588
	0.70	400	200	100	0.4259	0.3949	0.3501	0.3265	0.3203	0.3180	0.3200
	0.72	400	200	100	0.5821	0.5523	0.5313	0.5257	0.5257	0.5211	0.5255
	0.74	400	200	100	0.7329	0.7114	0.7092	0.7209	0.7287	0.7293	0.7342
	0.76	400	200	100	0.8412	0.8279	0.8442	0.8644	0.8757	0.8757	0.8794
	0.78	400	200	100	0.9202	0.9118	0.9349	0.9519	0.9607	0.9614	0.9627
	0.80	400	200	100	0.9710	0.9665	0.9806	0.9885	0.9907	0.9915	0.9916
	0.62	300	300	150	0.0609	0.0549	0.0325	0.0187	0.0119	0.0098	0.0086
	0.64	300	300	150	0.1290	0.1169	0.0812	0.0550	0.0387	0.0327	0.0301
	0.66	300	300	150	0.2485	0.2315	0.1829	0.1443	0.1221	0.1083	0.1030
	0.68	300	300	150	0.4025	0.3848	0.3363	0.3015	0.2710	0.2573	0.2520
	0.70	300	300	150	0.5935	0.5747	0.5433	0.5208	0.4994	0.4838	0.4806
	0.72	300	300	150	0.7551	0.7392	0.7255	0.7257	0.7158	0.7087	0.7073
	0.74	300	300	150	0.8760	0.8644	0.8661	0.8781	0.8800	0.8784	0.8787
	0.76	300	300	150	0.9554	0.9512	0.9566	0.9637	0.9668	0.9663	0.9657
	0.78	300	300	150	0.9840	0.9828	0.9863	0.9903	0.9932	0.9932	0.9936
	0.80	300	300	150	0.9966	0.9964	0.9975	0.9983	0.9989	0.9991	0.9990
	0.62	400	400	100	0.0522	0.0476	0.0172	0.0068	0.0048	0.0043	0.0046
	0.64	400	400	100	0.1056	0.0991	0.0494	0.0286	0.0200	0.0177	0.0192
	0.66	400	400	100	0.1963	0.1879	0.1233	0.0918	0.0813	0.0801	0.0846
	0.68	400	400	100	0.3278	0.3212	0.2558	0.2281	0.2217	0.2249	0.2324
	0.70	400	400	100	0.4926	0.4881	0.4519	0.4485	0.4594	0.4676	0.4780
	0.72	400	400	100	0.6524	0.6500	0.6583	0.6904	0.7159	0.7272	0.7384
	0.74	400	400	100	0.8054	0.8055	0.8351	0.8709	0.8955	0.9035	0.9093
	0.76	400	400	100	0.9045	0.9062	0.9385	0.9626	0.9764	0.9781	0.9805
	0.78	400	400	100	0.9573	0.9580	0.9811	0.9929	0.9964	0.9971	0.9969
	0.80	400	400	100	0.9882	0.9887	0.9965	0.9993	0.9996	0.9996	0.9996

Tables 6.8 and 6.9 show the results of the power simulations for $p_{2H} = 0.65 < p_2 = 0.70$ for the balanced $(n_1 = n_2)$ and unbalanced $(n_1 > n_2)$ case, respectively. In these cases, the Bayesian powers exceed the frequentist powers for almost every scenario where $a_{02} \geq 0.10$.

Table 6.8: Power of the Bayesian and frequentist method at different levels of p_1 and sample sizes (scenario 3: $p_{2H} = 0.65 < p_2 = 0.70, \Delta = 0.10$, balanced).

p_{2H}	p_1	n_{2H}	n_1	n_2	Frequentist	Bayesian method					
						$a_{02} = 0$	0.1	0.25	0.50	0.75	1.0
	0.62	100	100	100	0.0466	0.0353	0.0379	0.0464	0.0615	0.0792	0.0990
	0.64	100	100	100	0.0890	0.0668	0.0730	0.0873	0.1143	0.1369	0.1653
	0.66	100	100	100	0.1538	0.1183	0.1288	0.1531	0.1949	0.2254	0.2648
	0.68	100	100	100	0.2372	0.1918	0.2089	0.2426	0.3002	0.3463	0.3893
	0.70	100	100	100	0.3518	0.2976	0.3215	0.3669	0.4356	0.4859	0.5330
	0.72	100	100	100	0.4687	0.4068	0.4381	0.4945	0.5719	0.6260	0.6711
	0.74	100	100	100	0.6070	0.5476	0.5804	0.6430	0.7155	0.7599	0.7945
	0.76	100	100	100	0.7372	0.6802	0.7208	0.7742	0.8353	0.8711	0.8928
	0.78	100	100	100	0.8381	0.7907	0.8251	0.8715	0.9149	0.9366	0.9486
	0.80	100	100	100	0.9150	0.8859	0.9093	0.9381	0.9618	0.9726	0.9797
	0.62	200	200	200	0.0605	0.0493	0.0549	0.0672	0.0980	0.1289	0.1613
	0.64	200	200	200	0.1365	0.1151	0.1320	0.1628	0.2210	0.2648	0.3126
	0.66	200	200	200	0.2428	0.2142	0.2441	0.3003	0.3848	0.4522	0.5082
	0.68	200	200	200	0.4105	0.3734	0.4186	0.4837	0.5738	0.6384	0.6872
	0.70	200	200	200	0.5922	0.5520	0.6045	0.6754	0.7626	0.8158	0.8500
	0.72	200	200	200	0.7544	0.7207	0.7706	0.8284	0.8897	0.9193	0.9377
	0.74	200	200	200	0.8781	0.8556	0.8919	0.9281	0.9602	0.9748	0.9817
	0.76	200	200	200	0.9519	0.9410	0.9604	0.9777	0.9908	0.9954	0.9969
	0.78	200	200	200	0.9837	0.9793	0.9879	0.9941	0.9977	0.9987	0.9990
	0.80	200	200	200	0.9956	0.9949	0.9968	0.9986	0.9994	0.9996	0.9998
	0.62	400	200	200	0.0623	0.0519	0.0694	0.1015	0.1535	0.1995	0.2368
	0.64	400	200	200	0.1308	0.1111	0.1462	0.2123	0.3113	0.3770	0.4311
	0.66	400	200	200	0.2422	0.2141	0.2796	0.3771	0.4947	0.5689	0.6257
	0.68	400	200	200	0.4107	0.3711	0.4605	0.5770	0.7005	0.7625	0.8017
0.65	0.70	400	200	200	0.5848	0.5472	0.6512	0.7610	0.8518	0.8935	0.9171
	0.72	400	200	200	0.7513	0.7194	0.8104	0.8892	0.9444	0.9644	0.9744
	0.74	400	200	200	0.8766	0.8566	0.9181	0.9608	0.9846	0.9903	0.9942
	0.76	400	200	200	0.9526	0.9400	0.9744	0.9906	0.9966	0.9985	0.9993
	0.78	400	200	200	0.9846	0.9800	0.9941	0.9987	0.9997	1.0000	1.0000
	0.80	400	200	200	0.9967	0.9950	0.9987	0.9998	1.0000	1.0000	1.0000
	0.62	300	300	300	0.0767	0.0663	0.0770	0.1020	0.1493	0.1929	0.2351
	0.64	300	300	300	0.1879	0.1672	0.1972	0.2506	0.3299	0.3967	0.4604
	0.66	300	300	300	0.3519	0.3228	0.3727	0.4482	0.5524	0.6294	0.6851
	0.68	300	300	300	0.5667	0.5341	0.5923	0.6775	0.7716	0.8284	0.8675
	0.70	300	300	300	0.7618	0.7375	0.7913	0.8587	0.9159	0.9432	0.9591
	0.72	300	300	300	0.9054	0.8872	0.9215	0.9529	0.9775	0.9878	0.9910
	0.74	300	300	300	0.9724	0.9675	0.9785	0.9908	0.9970	0.9986	0.9994
	0.76	300	300	300	0.9920	0.9902	0.9949	0.9980	0.9997	0.9998	0.9999
	0.78	300	300	300	0.9986	0.9982	0.9990	0.9999	1.0000	1.0000	1.0000
	0.80	300	300	300	0.9999	0.9999	0.9999	1.0000	1.0000	1.0000	1.0000
	0.62	400	400	400	0.0847	0.0754	0.0924	0.1274	0.1879	0.2468	0.3031
	0.64	400	400	400	0.2300	0.2095	0.2544	0.3239	0.4275	0.5118	0.5759
	0.66	400	400	400	0.4489	0.4198	0.4870	0.5818	0.6954	0.7686	0.8142
	0.68	400	400	400	0.6905	0.6666	0.7316	0.8129	0.8856	0.9266	0.9476
	0.70	400	400	400	0.8732	0.8566	0.9013	0.9418	0.9731	0.9843	0.9892
	0.72	400	400	400	0.9636	0.9589	0.9759	0.9884	0.9958	0.9980	0.9993
	0.74	400	400	400	0.9925	0.9914	0.9959	0.9989	0.9996	0.9998	1.0000
	0.76	400	400	400	0.9994	0.9993	0.9996	0.9999	1.0000	1.0000	1.0000
	0.78	400	400	400	0.9999	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000
	0.80	400	400	400	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

6.5 Real data analysis and sample size simulations

6.5.1 Real data analysis

Eron et al. (2006) and Carosi et al. (2009) present the non-inferiority clinical trials in HIV patients conducted by the same sponsor. They used the same primary endpoint, which is the proportion of treatment success patients, defined as the proportion satisfying HIV-1 RNA concentrations of < 400 copies/mL at 48 weeks. The objective of the trial by Eron et al. (2006) is to show the non-inferior antiviral activity of fosamprenavir/ritonavir (FPV/r) 700mg/100mg twice daily (BID) relative to lopinavir/ritonavir (LPV/r) 400mg/100mg BID, each with the co-formulation of abacavir-lamivudine (ABC/3TC) 600mg/300mg once daily (QD). On the other hand, the objective of the trial by Carosi et al. (2009) is to show the non-inferior antiviral activity of FPV/r 1400mg/100mg QD relative to FPV/r 700mg/100mg BID, each with the co-formulation of ABC/3TC 600mg/300mg QD. In addition, the trial by *et al.* Eron et al. (2006) finished not so long before Carosi et al. (2009) began. Based on the six criteria by Pocock (1976), Cuffe (2011) stated that the two trials are similar. Then, we apply our proposed method by treating Carosi et al. (2009) as the present trial and Eron et al. (2006) as the historical trial. Here, we utilize the historical data of the FPV/r 700mg/100mg BID group in Eron et al. (2006) to construct the prior distribution of the same group of the present trial. In the following, let π_1 and π_2 be the treatment success probabilities of the FPV/r 1400mg/100mg QD group and the FPV/r 700mg/100mg BID group, respectively.

Table 6.10: Historical trial data.

	HIV-1 RNA-1 <400 copies/mL	HIV-1 RNA-1 ≥400 copies/mL	total
FPV/r 700mg/100mg BID	315 (72.6%)	119 (27.4%)	434
LPV/r 400mg/100mg BID	317 (71.4%)	127 (28.6%)	444

Table 6.10 shows the result of the historical trial. Non-inferiority of the FPV/r 700mg/100mg BID relative to LPV/r 400mg/100mg BID is confirmed based on the frequentist two sided Wald 95% CI $(-0.0484, 0.0705)$ with non-inferiority margin $\Delta = 0.12$ stated in Eron et al. (2006). Our main interest in this trial is, however, not the confirmation of non-inferiority, but the proportion of treatment success patients of the FPV/r 700mg/100mg BID group. Here, we construct the prior

distribution of the FPV/r 700mg/100mg BID group for the present trial. First, we construct the conditional power prior from the historical trial data in Table 6.10 and $f_{02}(\pi_2) \propto (1-\pi_2)^{-1}$ as follows:

$$\begin{aligned} f(\pi_2 \mid X_{2H}, a_{02}) &\propto L(\pi_2 \mid X_{2H})^{a_{02}} \cdot f_{02}(\pi_2) \\ &= \pi_2^{315a_{02}}(1-\pi_2)^{119a_{02}} \cdot (1-\pi_2)^{-1} \\ &= \pi_2^{(315a_{02}+1)-1}(1-\pi_2)^{119a_{02}-1}. \end{aligned}$$

Then, we let the prior of π_2 be $Beta([315a_{02}]+1, [119a_{02}])$. However, because we do not have any prior information for the FPV/r 1400mg/100mg QD group, we let the prior of π_1 be

$$f(\pi_1) \propto \pi_1^{-1}.$$

Table 6.11 shows the examples of the parameters of the prior distributions with several a_{02}'s. Here, ESS is the prior effective sample size defined in the study by Morita et al. (2008). It characterizes the information contained in the prior distributions. For the beta distribution $Beta(\alpha, \beta)$, ESS $= \alpha + \beta$.

Table 6.11: Prior distribution of the FPV/r 700mg/100mg BID group constructed from the historical trial.

a_{02}	α_2	β_2	ESS
0	(1)	(0)	(1)
0.01	4	1	5
0.025	8	2	10
0.1	32	11	43
0.25	79	29	108
0.5	158	59	217

Table 6.12 shows the result of the present trial. The two sided 95% Wald confidence interval for the difference $\pi_1 - \pi_2$ was $(-0.114, 0.095)$ from the frequentist perspective stated in Carosi et al. (2009). Therefore, non-inferiority was established with the non-inferiority margin $\Delta = 0.12$ without using the historical data. Then, we consider the case with $\Delta = 0.10$, which is the second

most frequently used margin in the recent HIV non-inferiority trials in Flandre (2011). In this case, non-inferiority could not be demonstrated from the frequentist perspective. Here, we compare the result of our Bayesian methods to both values of Δ.

Table 6.12: Present trial data.

	HIV-1 RNA-1 <400 copies/mL	HIV-1 RNA-1 ≥400 copies/mL	total
FPV/r 1400mg/100mg QD	87 (82.1%)	19 (17.9%)	106
FPV/r 700mg/100mg BID	86 (81.1%)	20 (18.9%)	106

Table 6.13 lists the posterior probabilities for $a_{02} = 0, 0.01, 0.025, 0.1, 0.25$ (we have not shown the result of $a_{02} = 0.5$ because ESS exceeded the sample size of the control group n_2). Here, we used the exact expression (6.1). When $a_{02} = 0$, non-inferiority was established with $\Delta = 0.12$ and not established with $\Delta = 0.10$. The results were identical to those of the frequentist method. On the other hand, when $a_{02} \geq 0.1$, non-inferiority was established even with $\Delta = 0.1$ from the Bayesian perspective. This showed the feasibility of our Bayesian non-inferiority test.

Table 6.13: Bayesian posterior probability of the present trial.

		Δ	
a_{02}	ESS	0.12	0.10
0	(1)	0.9879	0.9701
0.01	5	0.9894	0.9735
0.025	10	0.9901	0.9749
0.1	43	0.9969	0.9910
0.25	108	0.9994	0.9979

6.5.2 Sample size calculation

Next, we consider sample size calculation by utilizing the historical data to reduce the sample size of the new trial. In the present study Carosi et al. (2009), it is stated that "assuming a 72% success rate in both treatment arms, a total of 364 subjects per arm would provide 90% power to assess the non-inferiority of the once-daily regimen compared with the twice-daily regimen at the one-sided 0.025 level of significance. Non-inferiority was defined as the lower bound of the two sided 95$\%$ confidence interval (CI) for the treatment difference being above -12%". However, our calculations based on the following formula for the non-inferiority test for the hypotheses $H_0 : p_1 \leq p_2 - \Delta$ versus $H_1 : p_1 > p_2 - \Delta$ in Chow *et al.* Chow et al. (2007) with $\alpha = 0.025, p_1 = p_2 = 0.72, \Delta = 0.12$ (n is the number of subjects per arm for balanced allocation)

$$n = \frac{(z_\alpha + z_\beta)^2\{p_1(1-p_1) + p_2(1-p_2)\}}{(p_1 - p_2 + \Delta)^2},$$

show that a total of 295 subjects per arm would provide 90$\%$ power, where z_α is the lower $100 \cdot \alpha$ percentile of the standard normal distribution. Hence, we treat 295 subjects per arm (total 590 subjects) as the frequentist sample size in the above setting. We conduct the sample size simulation based on our method providing 90$\%$ power with $p_1 = p_2 = 0.72$. Here, we let the prior distributions of π_2 be the ones shown in Table 6.11, and that of π_1 be proportional to π_1^{-1}.

Simulation procedure:

1. Specify N, M, and adequately small (n_1, n_2). Set COUNT $= 0$.

2. From the historical data, construct the prior distribution of π_2 shown in Table 6.11, and let the prior distribution of π_1 be $f(\pi_1) \propto \pi_1^{-1}$.

3. Generate $X_1 \sim Bin(n_1, p_1), X_2 \sim Bin(n_2, p_2)$, independently. In the following, suppose $X_1 > 0$.

4. Generate $\pi_{1,i}, \pi_{2,i}$ for $i = 1, \ldots, M$ independently from the posterior distributions of π_1 and π_2:

$$\pi_{1,i} \sim Beta(a_1, b_1), \pi_{2,i} \sim Beta(a_2, b_2),$$

   where $a_1 := X_1, b_1 := (n_1 - X_1) + 1, a_2 := \alpha_2 + X_2, b_2 := \beta_2 + (n_2 - X_2)$.

5. Calculate the estimated posterior probability

$$\widehat{P}(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) = \frac{1}{M}\sum_{i=1}^{M} I(\pi_{i,1} > \pi_{i,2} - \Delta).$$

6. If $\widehat{P}(\pi_1 > \pi_2 - \Delta \mid X_1, X_2) > 0.975$, then increase the COUNT by 1; otherwise, by 0.

7. Go back to step 3. and repeat N times.

8. Calculate the power by COUNT $/N$.

9. Increase (n_1, n_2) and set COUNT $= 0$, go back to step 3., and repeat until the power exceeds 0.90.

Parameter settings:

- N=10,000, $M = 10^6$, $\Delta = 0.12$.
- $a_{02} = 0, 0.01, 0.025, 0.1, 0.25, 0.5$.
- $p_1 = p_2 = 0.72$.
- Scenarios:
  - (a) Balanced allocation.
    - $(n_1, n_2) = (200, 200), (210, 210), \ldots, (390, 390)$.
  - (b) $2:1$ allocation.
    - For $a_{02} < 0.5$, $(n_1, n_2) = (300, 150), (310, 155), \ldots, (490, 245)$.
    - For $a_{02} = 0.5$, $(n_1, n_2) = (200, 100), (210, 105), \ldots, (390, 195)$.
  - (c) $a_{02} = 0.1, 0.25$ and $n_1 \approx$ ESS $+ n_2$.
    - For $a_{02} = 0.1$, $(n_1, n_2) = (300, 250), (310, 260), \ldots, (400, 350)$. (ESS $= 43$)
    - For $a_{02} = 0.25$, $(n_1, n_2) = (250, 150), (260, 160), \ldots, (350, 250)$. (ESS $= 108$)
  - (d) $a_{02} = 0.5$ and $n_2 = 220$.
    - $(n_1, n_2) = (150, 220), (160, 220), \ldots, (300, 220)$. (ESS $= 217$)

Table 6.14 shows the obtained results. For $a_{02} \leq 0.025$ (ESS ≤ 10), sample sizes changed slightly and some changes would be caused by the Monte Carlo error. For $a_{02} = 0.1$ (ESS $= 43$), the total sample size was 590 from scenario (c) and was less than that for $a_{02} \leq 0.025$. However, the frequentist total sample size was 590. Therefore, in these situations, our method could not decrease the total sample size as compared to the frequentist method. On the other hand, when $a_{02} = 0.25$ (ESS $= 108$), total sample sizes were smaller than the frequentist one. Especially for scenario (b), it was 495. Therefore, our method could decrease 95 patients as compared to the frequentist method. When $a_{02} = 0.5$ (ESS $= 217$), the total sample size was 405 when $(n_1, n_2) = (270, 135)$. However, ESS, which could be interpreted as the sample size of the historical trial, was more than 1.5 times as n_2 (sample size of the present trial). It seems that this borrows too much information from the historical trial. Therefore, from scenario (d), i.e., $n_2 = 220 \approx$ ESS $= 217$, we consider that the total sample size 470 would be the minimum. Consequently, we could decrease 120 patients in this situation.

Table 6.14: Sample size at different scenarios and different levels of a_{02} ($p_1 = p_2 = p_{2H} = 0.72, \Delta = 0.12$).

		(a) balanced		(b) 2 : 1		(c) $a_{02} = 0.1, 0.25$, and $n_1 \approx$ ESS $+ n_2$		(d) $a_{02} = 0.5$, and $n_2 = 220$	
a_{02}	ESS	(n_1, n_2)	total	(n_1, n_2)	total	(n_1, n_2)	total	(n_1, n_2)	total
0	(1)	(320, 320)	640	(440, 220)	660	–	–	–	–
0.01	5	(310, 310)	620	(440, 220)	660	–	–	–	–
0.025	10	(320, 320)	640	(440, 220)	660	–	–	–	–
0.1	43	(300, 300)	600	(400, 200)	600	(320, 270)	590	–	–
0.25	108	(280, 280)	560	(330, 165)	495	(300, 200)	500	–	–
0.5	217	(250, 250)	500	(270, 135)	405	–	–	(250, 220)	470

Additionally, in order to investigate the minimum sample size for $a_{02} = 0.25$ (ESS $= 108$), we conducted simulations where n_1 are from 250 to 400 by increments of 10 and n_2 are from (ESS $\leq$)110 to 250 by increments of 10. Table 6.15 shows the result. The powers do not monotonically increase with n_1 and n_2. This might be because of the Monte Carlo error. Here, total sample size would be 490 when $(n_1, n_2) = (340, 150), (300, 190)$, 500 when $(n_1, n_2) = (350, 150)$, $(340, 160)$, $(330, 170)$, $(320, 180)$, $(300, 200)$, and 510 when $(n_1, n_2) = (400, 110)$, $(390, 120), \ldots, (300, 210)$ except $(320, 190)$.

Table 6.9: Power of the Bayesian and frequentist method at different levels of p_1 and sample sizes (scenario 3: $p_{2H} = 0.65 < p_2 = 0.70, \Delta = 0.10$, unbalanced).

p_{2H}	p_1	n_{2H}	n_1	n_2	Frequentist	Bayesian method					
						$a_{02} = 0$	0.1	0.25	0.50	0.75	1.0
	0.62	100	100	50	0.0393	0.0297	0.0309	0.0377	0.0612	0.0827	0.1102
	0.64	100	100	50	0.0657	0.0546	0.0593	0.0759	0.1154	0.1502	0.1846
	0.66	100	100	50	0.1058	0.0884	0.1022	0.1331	0.1940	0.2433	0.2900
	0.68	100	100	50	0.1626	0.1398	0.1619	0.2039	0.2863	0.3444	0.3986
	0.70	100	100	50	0.2284	0.2005	0.2367	0.3027	0.4117	0.4825	0.5430
	0.72	100	100	50	0.3204	0.2874	0.3376	0.4240	0.5463	0.6199	0.6745
	0.74	100	100	50	0.4250	0.3872	0.4554	0.5557	0.6799	0.7488	0.7925
	0.76	100	100	50	0.5409	0.5043	0.5739	0.6812	0.7919	0.8424	0.8785
	0.78	100	100	50	0.6593	0.6262	0.7009	0.8009	0.8840	0.9202	0.9403
	0.80	100	100	50	0.7610	0.7315	0.8055	0.8880	0.9454	0.9672	0.9766
	0.62	200	200	100	0.0511	0.0426	0.0507	0.0703	0.1150	0.1587	0.1997
	0.64	200	200	100	0.0950	0.0830	0.0999	0.1490	0.2317	0.3013	0.3607
	0.66	200	200	100	0.1779	0.1607	0.1965	0.2748	0.3914	0.4716	0.5385
	0.68	200	200	100	0.2853	0.2609	0.3225	0.4372	0.5766	0.6589	0.7156
	0.70	200	200	100	0.4318	0.4034	0.4905	0.6197	0.7513	0.8128	0.8582
	0.72	200	200	100	0.5734	0.5473	0.6485	0.7804	0.8820	0.9208	0.9423
	0.74	200	200	100	0.7205	0.6980	0.7930	0.8881	0.9486	0.9686	0.9791
	0.76	200	200	100	0.8430	0.8295	0.9014	0.9568	0.9867	0.9935	0.9960
	0.78	200	200	100	0.9227	0.9123	0.9593	0.9875	0.9962	0.9981	0.9992
	0.80	200	200	100	0.9692	0.9655	0.9884	0.9974	0.9994	0.9997	0.9997
	0.62	400	200	100	0.0453	0.0395	0.0578	0.1019	0.1784	0.2368	0.2793
	0.64	400	200	100	0.1020	0.0889	0.1369	0.2256	0.3422	0.4227	0.4755
	0.66	400	200	100	0.1816	0.1604	0.2488	0.3903	0.5450	0.6264	0.6745
	0.68	400	200	100	0.2851	0.2609	0.3997	0.5751	0.7305	0.7940	0.8321
	0.70	400	200	100	0.4233	0.3968	0.5791	0.7577	0.8734	0.9137	0.9336
0.65	0.72	400	200	100	0.5789	0.5525	0.7434	0.8905	0.9541	0.9713	0.9779
	0.74	400	200	100	0.7225	0.6986	0.8699	0.9570	0.9849	0.9925	0.9946
	0.76	400	200	100	0.8438	0.8280	0.9491	0.9886	0.9970	0.9982	0.9992
	0.78	400	200	100	0.9212	0.9120	0.9824	0.9981	0.9996	0.9998	0.9999
	0.80	400	200	100	0.9645	0.9596	0.9946	0.9999	1.0000	1.0000	1.0000
	0.62	300	300	150	0.0569	0.0503	0.0659	0.1000	0.1679	0.2282	0.2875
	0.64	300	300	150	0.1305	0.1211	0.1549	0.2257	0.3464	0.4358	0.5037
	0.66	300	300	150	0.2442	0.2285	0.2997	0.4135	0.5737	0.6670	0.7261
	0.68	300	300	150	0.4029	0.3820	0.4882	0.6250	0.7693	0.8412	0.8780
	0.70	300	300	150	0.5848	0.5652	0.6805	0.8106	0.9105	0.9445	0.9643
	0.72	300	300	150	0.7591	0.7436	0.8432	0.9288	0.9744	0.9864	0.9916
	0.74	300	300	150	0.8832	0.8733	0.9380	0.9795	0.9944	0.9976	0.9985
	0.76	300	300	150	0.9523	0.9482	0.9808	0.9966	0.9996	0.9997	0.9999
	0.78	300	300	150	0.9872	0.9851	0.9965	0.9995	0.9998	0.9999	0.9999
	0.80	300	300	150	0.9960	0.9955	0.9994	0.9999	1.0000	1.0000	1.0000
	0.62	400	400	100	0.0529	0.0488	0.0717	0.1367	0.2638	0.3582	0.4341
	0.64	400	400	100	0.1063	0.1009	0.1626	0.3022	0.4941	0.6079	0.6845
	0.66	400	400	100	0.2047	0.1976	0.3271	0.5399	0.7368	0.8237	0.8665
	0.68	400	400	100	0.3400	0.3323	0.5192	0.7491	0.8992	0.9404	0.9587
	0.70	400	400	100	0.4894	0.4844	0.7154	0.9000	0.9753	0.9889	0.9939
	0.72	400	400	100	0.6580	0.6563	0.8658	0.9703	0.9957	0.9991	0.9996
	0.74	400	400	100	0.7971	0.7981	0.9514	0.9954	0.9997	0.9998	0.9999
	0.76	400	400	100	0.9039	0.9051	0.9887	0.9996	0.9999	1.0000	1.0000
	0.78	400	400	100	0.9604	0.9615	0.9963	1.0000	1.0000	1.0000	1.0000
	0.80	400	400	100	0.9882	0.9887	0.9996	1.0000	1.0000	1.0000	1.0000

Table 6.15: Power of the Bayesian method at different levels of sample sizes ($p_1 = p_2 = p_{2H} = 0.72, \Delta = 0.12, a_{02} = 0.25$); bold characters indicate more than 90%.

	n_2														
n_1	110	120	130	140	150	160	170	180	190	200	210	220	230	240	250
250	0.8004	0.8233	0.8143	0.8293	0.8309	0.8434	0.8489	0.8529	0.8525	0.8605	0.8594	0.8660	0.8721	0.8708	0.8804
260	0.8117	0.8242	0.8173	0.8368	0.8391	0.8445	0.8547	0.8621	0.8620	0.8668	0.8741	0.8794	0.8836	0.8866	0.8807
270	0.8184	0.8312	0.8389	0.8452	0.8554	0.8587	0.8628	0.8638	0.8715	0.8765	0.8775	0.8861	0.8896	0.8985	0.8963
280	0.8341	0.8417	0.8454	0.8533	0.8603	0.8640	08713	0.8770	0.8818	0.8853	0.8834	0.8953	0.8972	0.8988	**0.9031**
290	0.8419	0.8505	0.8540	0.8666	0.8696	0.8870	0.8762	0.8821	0.8854	0.8928	0.8976	0.8982	0.8991	**0.9058**	**0.9119**
300	0.8482	0.8597	0.8613	0.8711	0.8768	0.8794	0.8944	0.8901	**0.9012**	**0.9005**	**0.9079**	**0.9095**	**0.9092**	**0.9129**	**0.9162**
310	0.8550	0.8682	0.8781	0.8781	0.8865	0.8844	0.8948	0.8909	0.8964	**0.9031**	**0.9110**	**0.9097**	**0.9187**	**0.9213**	**0.9252**
320	0.8571	0.8689	0.8780	0.8805	0.8859	0.8917	0.8957	**0.9065**	0.8983	**0.9120**	**0.9153**	**0.9162**	**0.9190**	**0.9217**	**0.9276**
330	0.8665	0.8785	0.8801	0.8925	0.8945	0.8950	**0.9050**	**0.9074**	**0.9115**	**0.9116**	**0.9237**	**0.9240**	**0.9260**	**0.9262**	**0.9310**
340	0.8733	0.8806	0.8921	0.8986	**0.9038**	**0.9066**	**0.9088**	**0.9136**	**0.9198**	**0.9229**	**0.9261**	**0.9268**	**0.9291**	**0.9349**	**0.9409**
350	0.8842	0.8935	0.8950	0.8998	**0.9080**	**0.9097**	**0.9129**	**0.9234**	**0.9250**	**0.9226**	**0.9291**	**0.9339**	**0.9381**	**0.9372**	**0.9415**
360	0.8855	0.8934	0.8987	0.8970	**0.9096**	**0.9146**	**0.9166**	**0.9230**	**0.9255**	**0.9303**	**0.9350**	**0.9337**	**0.9453**	**0.9413**	**0.9443**
370	0.8928	0.8966	**0.9033**	**0.9066**	**0.9104**	**0.9183**	**0.9206**	**0.9306**	**0.9295**	**0.9363**	**0.9371**	**0.9415**	**0.9447**	**0.9519**	**0.9449**
380	0.8956	**0.9034**	**0.9056**	**0.9166**	**0.9147**	**0.9246**	**0.9295**	**0.9281**	**0.9339**	**0.9416**	**0.9435**	**0.9468**	**0.9451**	**0.9502**	**0.9491**
390	0.8969	**0.9029**	**0.9140**	**0.9233**	**0.9257**	**0.9236**	**0.9287**	**0.9305**	**0.9404**	**0.9434**	**0.9442**	**0.9483**	**0.9493**	**0.9486**	**0.9579**
400	**0.9026**	**0.9068**	**0.9153**	**0.9232**	**0.9250**	**0.9285**	**0.9362**	**0.9440**	**0.9402**	**0.9466**	**0.9474**	**0.9485**	**0.9556**	**0.9561**	**0.9555**

Therefore, we could decrease the total sample size to 490. Moreover, if we want to increase the sample size of the trial treatment group n_1 with the fixed total sample size, for example, 510, we can achieve more than 90$\%$ power when $(n_1, n_2) = (400, 110)$ for ESS $= 108$.

6.6 Discussion

In this chapter, we extended the fully Bayesian method in Gamalo et al. (2011) with pre-specified margin, and constructed the Bayesian framework which can be interpreted as an extension of Fisher's exact test. We showed the limiting relationship between Pr$(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)$ and the one sided p-value of Fisher's exact test, and derived the exact expression for the posterior probability Pr$(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)$ under mild conditions. We also evaluated the accuracy of the Monte Carlo approximation for the calculation of Pr$(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)$, and the operating characteristics for the Monte Carlo simulations. The results showed that the type I error rates were almost always controlled when the distribution of the historical data of the active control group was the same as that of the present data. From the simulations for the power and real data analysis, we showed that our methods can realize more efficient decision making when compared to the frequentist method with suitable historical data and parameter a_{02}. In the real data analysis, with adequate historical data and non-inferiority margin, our method demonstrated non-inferiority while the frequentist method could not. Further, by sample size calculation with adequate amount of historical data, the sample size could be reduced using our method. Further, we could flexibly change the allocation proportion with the (not minimum but nearly minimum) fixed sample size.

In the real data analysis and sample size calculations, we mainly focused on the situation where $n_2 \geq$ ESS in order to ensure that too much information was not borrowed from the historical trial. However, for example, for cases where the historical and present trial were almost identical, emergency situations, clinical trials for rare diseases, or pediatric trials borrowing strength from previous adult trials, this condition could be relaxed.

It should be noted that if the success probability in the historical data is lower than in the present data, type I error rates are inflated. Conversely, if the success probability in the historical data was higher than that in the present data, type I error rates were over-conservative. Similar tendencies were also reported in Gamalo et al. (2011) and Gamalo-Siebers et al. (2016), which could be naturally interpreted. In these situations, we may change the threshold for regarding $\pi_1 > \pi_2 - \Delta$ as true (p^*) from 0.975 to another value. However, when p^* was changed, the limiting relationship between our Bayesian decision making and the one sided Fisher's exact test with significance level 2.5$\%$ also changed.

In future work, we will investigate the selection of suitable historical data with more mild conditions than the Pocock criteria, in order to make better decision criteria when the present data and the prior distributions are in significant conflict, to construct the prior distribution from two or more historical trials, and calculate the exact probability in a shorter time.

Chapter 7

Concluding remarks

In this chapter, we summarize this thesis, and present some ideas for the future studies.

In Chapter 2, we summarized the literature of the Bayesian decision making mainly focused on the posterior probabilities of some hypotheses being true. In Chapter 3, we evaluated the posterior probabilities of the superiority hypotheses being true for Poisson rate parameters λ_1, λ_2. First, we expressed the posterior probability Pr$(\lambda_1 < \lambda_2 \mid X_1, X_2)$ as the cumulative distribution function of well known distributions. Next, we investigate the relationship between the Bayesian posterior probability and the p-value of the frequentist conditional test. Then we considered the probability Pr$(\lambda_1/\lambda_2 < c \mid X_1, X_2)$ which corresponds to a more generalized hypothesis $H_1 : \lambda_1/\lambda_2 < c$. In Chapter 4, we evaluated the posterior probability of non-inferiority hypothesis being true Pr$(\lambda_1 < \lambda_2 + \Delta \mid X_1, X_2)$ under mild conditions. First, we derived a simple expression, and then we considered a natural framework with switching from non-inferiority test to superiority which, under some conditions, corresponds to the conditional test discussed in Chapter 3. In Chapter 5, we first evaluated the posterior probability of superiority hypothesis being true for the variances of normal distributions Pr$(\sigma_1^2 > \sigma_2^2 \mid \mathbf{x}_1, \mathbf{x}_2)$ for several situations. We derived quite simple expressions and showed the relationship to the p-value of the frequentist F-test. Next, we considered the Bayesian posterior probability of equivalence Pr$(1/\Delta < \sigma_1/\sigma_2 < \Delta \mid \mathbf{x}_1, \mathbf{x}_2)$ and also derived quite simple expressions. In Chapter 6, we derived the exact expression for the posterior probability for the non-inferiority hypothesis being true for the binomial probability Pr$(\pi_1 > \pi_2 - \Delta \mid X_1, X_2)$ under mild conditions, and derived a framework which can be considered as the Bayesian non-inferiority extension of Fisher's exact test.

In this thesis, we considered decision making in quite simple situations. For future work, we will investigate more complicated situations, such as where the parameters are affected by some covariates. Furthermore, in order to apply Bayesian decision making, how to choose suitable prior distributions is the crucial problem. In this thesis, we applied the conditional power prior for fixed parameters. However, as discussed in Chapter 6, when the distributions of the data of historical and present trials differ, type I and type II error rates would be strongly affected. In order to prevent such "prior-data conflict", for example, Gravestock and Held (2017) estimated parameters of the prior distributions by empirical Bayes methods. Including this approach, we need to investigate how to choose the suitable prior. Further, we considered utilizing only the data of one historical trial. We should investigate constructing the prior when the data of more than two historical trials can be utilized.

Sample size calculation is also an important problem. We discussed this for one situation in Chapter 6, but the sample size is also affected by the prior-data conflict. We have to consider appropriate methods which take the prior-data conflict into consideration.

Bibliography

Abramowitz, M. and Stegun, I. A. (1964). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation.

Altham, P. M. (1969). Exact Bayesian analysis of a 2×2 contingency table, and Fisher's "exact" significance test. *Journal of the Royal Statistical Society. Series B (Methodological)*, 31(2):261–269.

Arends, L. R., Hoes, A. W., Lubsen, J., Grobbee, D. E., and Stijnen, T. (2000). Baseline risk as predictor of treatment benefit: three clinical meta-re-analyses. *Statistics in Medicine*, 19(24):3497–3518.

Berger, J. O. (2013). *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media.

Berger, J. O., Pericchi, L. R., Ghosh, J., Samanta, T., De Santis, F., Berger, J., and Pericchi, L. (2001). Objective Bayesian methods for model selection: introduction and comparison. *Lecture Notes-Monograph Series*, pages 135–207.

Bosch, M. E. V. and Wall, M. (1997). Visual acuity scored by the letter-by-letter or probit methods has lower retest variability than the line assignment method. *Eye*, 11(3):411–417.

Carosi, G., Lazzarin, A., Stellbrink, H., Moyle, G., Rugina, S., Staszewski, S., Givens, N., Ross, L., Granier, C., Ait-Khaled, M., et al. (2009). Study of once-daily versus twice-daily fosamprenavir plus ritonavir administered with abacavir/lamivudine once daily in antiretroviral-naïve HIV-1–infected adult subjects. *HIV clinical trials*, 10(6):356–367.

Casella, G. and Berger, R. L. (1987). Reconciling Bayesian and frequentist evidence in the one-sided testing problem. *Journal of the American Statistical Association*, 82(397):106–111.

Chow, S.-C., Shao, J., and Wang, H. (2007). *Sample size calculations in clinical research*. CRC press.

Cohen, J. A., Coles, A. J., Arnold, D. L., Confavreux, C., Fox, E. J., Hartung, H.-P., Havrdova, E., Selmaj, K. W., Weiner, H. L., Fisher, E., et al. (2012). Alemtuzumab versus interferon beta 1a as first-line treatment for patients with relapsing-remitting multiple sclerosis: a randomised controlled phase 3 trial. *The Lancet*, 380(9856):1819–1828.

Committee for Proprietary Medicinal Products (CPMP) and others (2000). Points to consider on switching between superiority and non-inferiority. 2000.

Cuffe, R. L. (2011). The inclusion of historical control data may reduce the power of a confirmatory study. *Statistics in Medicine*, 30(12):1329–1338.

Doi, M. (2016). Bayesian index of superiority and the p-value of the conditional test for Poisson parameters. *Journal of the Japan Statistical Society*, 46(2):99–127.

Doi, M., Ide, K., and Kawasaki, Y. (2017a). Bayesian indexes of superiority and equivalence and the p-value of the F-test for the variances of normal distributions. *Japanese Journal of Biometrics*, 38(1):1–16.

Doi, M., Takahashi, F., and Kawasaki, Y. (2017b). Bayesian noninferiority test for 2 binomial probabilities as the extension of Fisher exact test. *Statistics in Medicine*, 36(30):4789–4803.

Eron, J., Yeni, P., Gathe, J., Estrada, V., DeJesus, E., Staszewski, S., Lackey, P., Katlama, C., Young, B., Yau, L., et al. (2006). The klean study of fosamprenavir-ritonavir versus lopinavir-ritonavir, each in combination with abacavir-lamivudine, for initial treatment of HIV infection over 48 weeks: a randomised non-inferiority trial. *The Lancet*, 368(9534):476–482.

Flandre, P. (2011). Statistical methods in recent HIV noninferiority trials: reanalysis of 11 trials. *PloS one*, 6(9):e22871.

Gamalo, M. A., Tiwari, R. C., and LaVange, L. M. (2014). Bayesian approach to the design and analysis of non-inferiority trials for anti-infective products. *Pharmaceutical Statistics*, 13(1):25–40.

Gamalo, M. A., Wu, R., and Tiwari, R. C. (2011). Bayesian approach to noninferiority trials for proportions. *Journal of Biopharmaceutical Statistics*, 21(5):902–919.

Gamalo, M. A., Wu, R., and Tiwari, R. C. (2016). Bayesian approach to non-inferiority trials for normal means. *Statistical Methods in Medical Research*, 25(1):221–240.

Gamalo-Siebers, M., Gao, A., Lakshminarayanan, M., Liu, G., Natanegara, F., Railkar, R., Schmidli, H., and Song, G. (2016). Bayesian methods for the design and analysis of noninferiority trials. *Journal of Biopharmaceutical Statistics*, 26(5):823–41.

Ghosh, J., Purkayastha, S., and Samanta, T. (2005). Role of p-values and other measures of evidence in Bayesian analysis. *Handbook of Statistics*, 25:151–170.

Ghosh, P., Nathoo, F., Gönen, M., and Tiwari, R. C. (2011). Assessing noninferiority in a three-arm trial using the Bayesian approach. *Statistics in Medicine*, 30(15):1795–1808.

Ghosh, S., Ghosh, S., and Tiwari, R. C. (2016). Bayesian approach for assessing non-inferiority in a three-arm trial with pre-specified margin. *Statistics in Medicine*, 35(5):695–708.

Glass, G. V., Peckham, P. D., and Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of educational research*, 42(3):237–288.

Gould, A. L. (1991). Using prior findings to augment active-controlled trials and trials with small placebo groups. *Drug Information Journal*, 25(3):369–380.

Gravestock, I. and Held, L. (2017). Adaptive power priors with empirical Bayes for clinical trials. *Pharmaceutical Statistics*.

Gu, K., Ng, H. K. T., Tang, M. L., and Schucany, W. R. (2008). Testing the ratio of two Poisson rates. *Biometrical Journal*, 50(2):283–298.

Howard, J. (1998). The 2× 2 table: A discussion from a Bayesian viewpoint. *Statistical Science*, 13(4):351–367.

Ibrahim, J. G. and Chen, M.-H. (2000). Power prior distributions for regression models. *Statistical Science*, 15(1):46–60.

ICH E9 Expert Working Group (1999). ICH Harmonised Tripartite Guideline: Statistical Principles for Clinical Trials. *Statistics in Medicine*, 18(15):1905–1942.

Jeffreys, H. (1961). *The theory of probability*. OUP Oxford.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795.

Kawasaki, Y., Doi, M., Ide, K., and Miyaoka, E. (2016). A Bayesian non-inferiority index for two Poisson parameters. *Cogent Mathematics*, 3(1):1229373.

Kawasaki, Y. and Miyaoka, E. (2010). A posterior density for the difference between two binominal proportions and the highest posterior density credible interval. *Journal of the Japan Statistical Society*, 40(2):265–275.

Kawasaki, Y. and Miyaoka, E. (2012a). A Bayesian inference of $P(\lambda_1 < \lambda_2)$ for two Poisson parameters. *Journal of Applied Statistics*, 39(10):2141–2152.

Kawasaki, Y. and Miyaoka, E. (2012b). A Bayesian inference of $P(\pi_1 > \pi_2)$ for two proportions. *Journal of Biopharmaceutical Statistics*, 22(3):425–437.

Kawasaki, Y. and Miyaoka, E. (2013). A Bayesian non-inferiority test for two independent binomial proportions. *Pharmaceutical Statistics*, 12(4):201–206.

Kawasaki, Y. and Miyaoka, E. (2014). Comparison of three calculation methods for a Bayesian inference of two Poisson parameters. *Journal of Modern Applied Statistical Methods*, 13(1):26.

Kawasaki, Y., Shimokawa, A., and Miyaoka, E. (2014). On the Bayesian index of superiority and the p-value of the Fisher exact test for binomial proportions. *Journal of the Japan Statistical Society*, 44(1):73–81.

Krishnamoorthy, K. and Thomson, J. (2004). A more powerful test for comparing two Poisson means. *Journal of Statistical Planning and Inference*, 119(1):23–35.

Markowski, C. A. and Markowski, E. P. (1990). Conditions for the effectiveness of a preliminary test of variance. *The American Statistician*, 44(4):322–326.

Massacesi, L., Tramacere, I., Amoroso, S., Battaglia, M. A., Benedetti, M. D., Filippini, G., La Mantia, L., Repice, A., Solari, A., Tedeschi, G., et al. (2014). Azathioprine versus beta interferons for relapsing-remitting multiple sclerosis: a multicentre randomized non-inferiority trial. *PLoS One*, 9(11):e113371.

Morita, S., Thall, P. F., and Müller, P. (2008). Determining the effective sample size of a parametric prior. *Biometrics*, 64(2):595–602.

Pericchi, L. R. (2005). Model selection and hypothesis testing based on objective probabilities and Bayes factors. *Handbook of Statistics*, 25:115–149.

Pocock, S. J. (1976). The combination of randomized and historical controls in clinical trials. *Journal of chronic diseases*, 29(3):175–188.

Price, R. M. and Bonett, D. G. (2000). Estimating the ratio of two Poisson rates. *Computational Statistics & Data Analysis*, 34(3):345–356.

Przyborowski, J. and Wilenski, H. (1940). Homogeneity of results in testing samples from Poisson series: With an application to testing clover seed for dodder. *Biometrika*, 31(3/4):313–323.

Robert, C. (2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Science & Business Media.

Rosser, D. A., Murdoch, I. E., and Cousens, S. N. (2004). The effect of optical defocus on the test–retest variability of visual acuity measurements. *Investigative ophthalmology & visual science*, 45(4):1076–1079.

Rothman, K. J., Greenland, S., and Lash, T. L. (2008). *Modern epidemiology*. Lippincott Williams & Wilkins.

US Food and Drug Administration and others (2010). Guidance for the use of Bayesian statistics in medical device clinical trials. *Silver Spring, Maryland: FDA*.

US Food and Drug Administration and others (2016). Non-inferiority clinical trials to establish effectiveness. *Silver Spring, Maryland: FDA*.

Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika*, 29(3/4):350–362.

Wilcox, R. R. (1995). ANOVA: A paradigm for low power and misleading measures of effect size? *Review of Educational Research*, 65(1):51–77.

Zaslavsky, B. G. (2010). Bayesian versus frequentist hypotheses testing in clinical trials with dichotomous and countable outcomes. *Journal of Biopharmaceutical Statistics*, 20(5):985–997.

Zaslavsky, B. G. (2013). Bayesian hypothesis testing in two-arm trials with dichotomous outcomes. *Biometrics*, 69(1):157–163.

Zwillinger, D. (2014). *Table of integrals, series, and products*. Elsevier.